

مبادئ الفكر الاقتصادي

تأليف
الدكتور عبد العزيز فهمي قبيص

استاذ الاحصاء
بالمركز الدولي لتعليم الاحصاء
في بيروت

الطبعة الاولى

١٩٦٦

حقوق الطبع محفوظة للمؤلف

بيروت - لبنان

مبادئ الأساليب الاحصائية

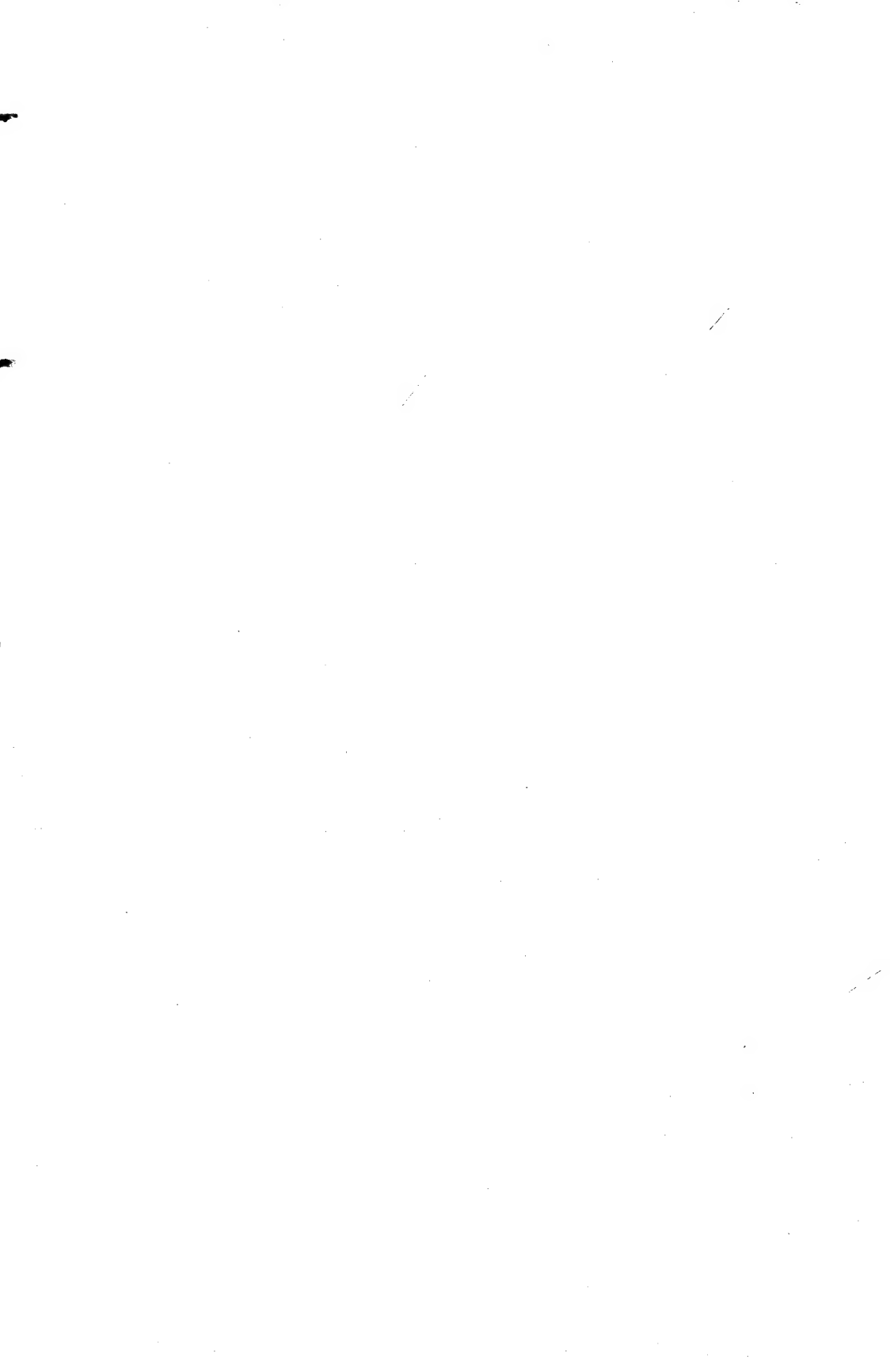
مُقَدِّمَة

أقدم هذا الكتاب الى الطالب المبتدىء في دراسة علم الإحصاء ، وقد راعيت فيه أن أناقش بشيء من التفصيل النواحي العملية المختلفة التي تسبق التحليل الرياضي للبيانات الاحصائية ، حيث اني لاحظت خلال السنوات الطويلة التي قمت فيها بتدريس هذا الموضوع ان الطالب بسبب تركيز كتب الاحصاء على الناحية الرياضية لا يدرك أهمية النواحي العملية التي يتطلبها البحث الاحصائي . ولا شك ان عدم إدراك الطالب لأهمية هذه النواحي يُبعد عن ذهنه الهدف الأساسي من تدريسه علم الاحصاء. ان الهدف الاسامي من تدريس طالب التجارة او الزراعة او الاجتماع علم الاحصاء هو تزويده بأسلوب في البحث العلمي يستطيع ان يعتمد عليه كلما أراد بحث مشكلة من المشاكل بحثاً علمياً صحيحاً . لعل في هذه المناقشة الموجزة للنواحي العملية للأسلوب الاحصائي في البحث العلمي ما يشجع الطالب على محاولة التوسع فيها والاستزادة منها بالرجوع الى مراجعها في اللغات الاجنبية المختلفة . ذلك لأن الدول العربية المختلفة في هذه المرحلة من تاريخها هي في أشد الحاجة الى جيل من المتخصصين في البحوث العلمية حتى تهتدي بهم في تخطيطها لشؤونها الاقتصادية والاجتماعية والعمرانية .

واني مدين بالشكر الوافر للاستاذ مصطفى كريسديه مدير دار النهضة العربية الذي تحمل عني عبء طبع هذا الكتاب وإخراجه على احسن وجه .

بيروت في اول آب ١٩٦٦

عبد العزيز فهمي هيكل



الفصل الأول

طبيعة علم الاحصاء

كان الانسان ولا زال يعمل جاهداً في البحث عن حقائق الكون المحيط به وأسرار الحياة التي يحلها . وكان في بادىء الامر يعتمد في بحثه هذا على تأملاته الخاصة فتوصل بذلك إلى فلسفات مختلفة متضاربة كان يتأرجح بالنسبة لها بين اليقين قارة والشك قارة اخرى . ولا يجب ان نقلل من أهمية الافكار التي تضمنتها هذه الفلسفات بالرغم من تضاربها ، إذ انها أخرجت الانسان من حالة المعجز الفكري السقي كانت تسيطر عليه ومهدت بذلك الطريق الى البحث العلمي .

ولا شك ان الانسان لم يقتنع بالنتائج التي توصل اليها عن طريق تأملاته حيث لم يستطع أن يسبغ عليها صفة الحقيقة المجردة ، وبذلك بدأ يفكر في منهاج آخر يعتمد عليه في بحثه . ان الملاحظة العابرة لا يمكن أن تعتبر حقيقة علمية بالمعنى الصحيح مهما بلغت أهميتها إلا إذا ثبت بالبرهان المادي الملموس ما لا يدعو الى الشك فيها ، وكيف السبيل الى إثباتها بغير التجربة في العمل؟

وبذلك بدأ الإنسان يعتمد على التجربة في العمل كمنهاج لبحثه عن الحقيقة . ولا يعني ذلك أن فقدت ملاحظة الانسان أهميتها ، إذ ظلت هي نقطة الانطلاق في كثير من الابحاث العلمية . ان ملاحظة الباحث لظاهرة وما

يتولد عن ذلك من إحساس بمشكلة تتطلب الحل او التفسير هي في الحقيقة الخطوة الاولى في البحث العلمي .

ولقد استطاع الانسان بمناهجه الجديد - التجربة في العمل - أن يتوصل الى كنز هائل من الحقائق العلمية في مختلف مجالات العلوم الطبيعية، الأمر الذي ساعد على تقدم هذه العلوم وعلى صلابة الأسس التي تعتمد عليها . وقد أفاد الانسان كثيراً من هذا التقدم حيث أصبحت حياته اليومية تقوم بالدرجة الأولى على استخدام مختلف الأشياء التي توصل إلى صنعها عن طريق البحث العلمي .

وبالرغم من تقدم العلوم الطبيعية تقدماً ملحوظاً بقيت العلوم الاجتماعية متخلفة وراءها حيث ان مجال استخدام التجارب في العمل في هذه العلوم مجال ضيق محدود، إذ من الصعب في كثير من الحالات أن نخضع سلوك الانسان لتجارب تجريها عليه في العمل حيث يقف دون ذلك عقبات كثيرة ترتبط بعادات المجتمع وتقاليده والأنظمة القانونية التي يسير عليها . على ان التقدم الحضاري في كثير من الدول وخاصة بعد الحرب العالمية الثانية قضى على كثير من هذه العقبات وبذلك بدأ العلماء يقومون بأبحاث على سلوك الانسان مستخدمين في ذلك منهج التجربة في العمل ، وبالرغم من ذلك لا نستطيع أن نتفائل كثيراً في امكانية اتساع نطاق استخدام هذا المنهج في ميدان العلوم الاجتماعية ، حيث لا توجد قواعد محددة ثابتة تحكم سلوك الانسان في كل مجتمع وفي كل زمن .

لم يقف الانسان مكتوف اليدين امام هذه المشكلة، إذ استطاع أن يهتدي إلى منهاج آخر يساعده في الكشف عن الحقائق الخاصة بسلوك الانسان في النواحي الاجتماعية والاقتصادية والسيكولوجية والتربوية، انه المنهج الاحصائي . وبذلك عندما ندرس الاحصاء انما ندرس في الواقع منهجاً من مناهج البحث العلمي . ولا يعني ذلك أن الاحصاء ليس علماً قائماً بذاته، فهو في الواقع علم له

قوانينه وقواعده الرياضية الخاصة به ولكن مجال تطبيقه هو في خدمة العلوم الأخرى .

على أن استخدام المنهاج الاحصائي في البحث العلمي لا يقتصر فقط على العلوم الاجتماعية ، بل يمتد ايضاً الى بعض النواحي في العلوم الطبيعية ، وبشكل عام نستطيع أن نقول انه في كل حالة لا تتوفر بالنسبة لها المعلومات الكاملة عن الظواهر موضوع البحث يمدنا المنهاج الاحصائي (المعاينة) بالقواعد والاساليب التي يمكن استخدامها للتوصل الى قرارات حكيمة يمكن أن نطمئن اليها . فإذا أردنا مثلاً أن نقرر اذا كان نوع معين من السماد أفضل من غيره لا يستطيع الباحث أن يقوم بالتجارب على جميع قطع الأرض الموجودة في الدولة وبذلك يضطر الى أخذ عينة من الأراضي لاجراء تجاربه ، الأمر الذي يجعله في حاجة إلى فهم أسلوب المعاينة حتى يكون بحثه علمياً صحيحاً . كذلك إذا أردنا أن نفهم انماط الاستهلاك في مجتمع معين لا يستطيع الباحث أن يقوم بجمع المعلومات من جميع أسر المجتمع ، فيكون بذلك مضطراً الى أخذ عينة من الاسر ، كيف يستطيع اختيار هذه الأسر؟ ، كيف يمكن أن تكون العينة غير متحيزة ؟ ، كيف يمكن أن يقدر منها المقاييس العامة لانماط الاستهلاك في المجتمع ؟ ، كيف يستطيع أن يحسب أخطاء هذه التقديرات ؟ ، كل هذه الاسئلة وغيرها لا يمكن الاجابة عليها الا بفهم أسلوب المعاينة .

ولا يجب أن يتبادر الى الذهن ان المنهاج الاحصائي يمكن استخدامه في بحث أية ظاهرة مهما كان نوعها ، ذلك لأن هذا المنهاج يبدأ أولاً بجمع المعلومات عن الظاهرة أو الظواهر موضوع البحث ، فإذا لم تكن هذه المعلومات هي نفسها عبارة عن ارقام أو يمكن تحويلها الى ارقام يتعذر بذلك استخدام المنهاج الاحصائي . الا ان ذلك لا يجد كثيراً من مجالات استخدام المنهاج الاحصائي إذ يستطيع الباحث أن يصيغ الاسئلة التي يجمع

بها المعلومات في شكل يساعده على اتباع هذا المنهاج ؛ فاذا أردنا مثلا أن نبحث رأي المجتمع في سياسة الحكومة نحو التعليم وسألنا كل مستجوب عن رأيه في هذه السياسة فالتنا بهذا السؤال سوف نحصل على إجابات يستحيل تحويلها إلى معلومات رقمية تساعدنا في اتباع المنهاج الاحصائي ، أما إذا كانت الأسئلة في صيغة أخرى ، مثلا : هل توافق على سياسة الحكومة نحو تحديد سن الالتحاق بالمداس الابتدائية ؟ وجعلنا الاجابة بنعم أم لا ، فالتنا بمثل هذا السؤال نستطيع أن نحصل على اجابات يمكن تحويلها الى أرقام حيث نستطيع أن نحدد عدد الموافقين على هذه السياسة وعدد غير الموافقين .

والاحصاء كمنهاج للبحث العلمي يساعد الانسان في اتخاذ قرارات حكيمة عند مواجهة عدم التأكد ولذلك ازدادت أهميته في العصر الحاضر زيادة كبيرة ، حيث نواجه عدم التأكد في حياتنا اليومية وفي حياتنا العامة بشكل لم يسبق له مثيل . اننا نناقش كل ما توارثناه عن آباءنا وأجدادنا من أفكار وتعاليم ونطلب إجابة علمية على كل ما نوجه من أسئلة . ان أسئلة مثل : هل هناك علاقة بين التدخين والسرطان ؟ وهل هناك علاقة بين تغذية الطفل وذكاءه؟ وهل هناك اختلاف جوهري بين خصوبة المرأة المتعلمة وغير المتعلمة ؟ وهل فعلا يتناقص الميل الحدي للاستهلاك للفرد كلما زاد دخله ؟ وهل توجد فعلا علاقة عكسية بين سعر الفائدة والاستثمار ؟ وهل تتغير أنماط الاستهلاك بتغير مستويات الدخل وما نوع هذا التغير وكيف يمكن قياسه ؟ وهل تتناقص التكلفة الحدية لأية سلعة في بادىء الأمر عند زيادة الانتاج ثم تتزايد بعد ذلك ؟ ... الخ . اننا نقرأ عن اجابات لمثل هذه الأسئلة في الكتب المتخصصة ولكنها في الغالب تكون اجابات بمثابة لأراء أصحابها واعتقاداتهم . الأمر الذي يفقدها الصفة العلمية ويجعلها بذلك غير قادرة على اقناعنا . وحتى نستطيع الاجابة على هذه الاسئلة وغيرها اجابة علمية صحيحة يجب ان نتبع المنهاج الاحصائي في بحثنا .

وبذلك أصبح المنهاج الاحصائي هو المنهاج الذي بغيره لا يقتنع إنسان العصر الحديث عندما يوجه سؤالاً يتطلب الاجابة العلمية الصحيحة عليه . ان رجل الأعمال لم يعد يثق في نتائج أية سياسة يتبناها نحو تسويق بضاعته إذا لم يكن المنهاج الاحصائي مؤيداً لها ، كما ان التخطيط الاقتصادي والاجتماعي لا يمكن أن يكون تخطيطاً علمياً صحيحاً ما لم تؤيده البيانات الاحصائية وما لم يستخدم المنهاج الاحصائي في تتبع تنفيذه وتقييم نتائجه .

انواع البحوث الاحصائية :

تبين لنا مما تقدم أن مجالات استخدام المنهاج الاحصائي واسعة جداً حيث يستخدم في جميع العلوم الاجتماعية منها والطبيعية وحيث يساعدنا في اتخاذ قرارات حكيمة عند مواجهة عدم التأكد ، الا اننا نستطيع أن نفرق بين ثلاث أنواع من البحوث الاحصائية هي :

أولاً . البحوث الاحصائية الوصفية وفيها تجمع المعلومات عن ظاهرة أو ظواهر معينة لا لخدمة هدف بذاته محدد سلفاً ، وانما بقصد توفير البيانات التي يمكن أن تخدم أغراضاً متعددة لباحثين مختلفين فيما بعد ، وفي الغالب تقوم الأجهزة الاحصائية العامة في الدولة بهذه البحوث الوصفية اما على فترات دورية كاهو الحال في تعدادات السكان والتعدادات الصناعية والزراعية والتجارية ، أو على فترات غير دورية كما هو الحال في بحث ميزانية الاسرة . فإذا أخذنا تعداد السكان مثلاً نلاحظ أن الجهاز الاحصائي المكلف به يقوم بجمع معلومات تمثل النواحي الديموجرافية والاجتماعية والاقتصادية للسكان وبذلك تكون مادة خام لاجراء البحوث التحليلية عليها من قبل الباحثين المهتمين بأية ناحية من هذه النواحي المختلفة . وفي التعداد الصناعي تجمع المعلومات عن القوة العاملة المستخدمة في الصناعة وعن الاجور المدفوعة لها وعن القيمة الصافية المضافة وعن الاستثمار الكلي والصافي في النشاط الصناعي

وغير ذلك من المعلومات التي تساعد في بحث أي موضوع من هذه الموضوعات بحثاً تحليلياً .

ثانياً : البحوث الاحصائية التحليلية - وهي التي تجمع فيها المعلومات التي تخدم هدف معين في ذهن الباحث أو التي تساعد في تفسير مشكلة معينة لاحظها الباحث أو لاختبار صحة فرض معين ؛ والباحث في هذا النوع من البحوث لا يقتصر فقط على جمع المعلومات التي تخدم هدفه وإنما يقوم كذلك بتحليلها التحليل الذي يراه مناسباً لاستنتاج المقاييس والمعاملات التي يرغب في التوصل إليها . وعند ما نتكلم عن جمع المعلومات في هذا النوع من البحوث لا نقصد أن الباحث لا بد أن يقوم هو نفسه بدراسة استقصائية في الميدان لجمع المعلومات ، وقد يضطر فعلاً الى ذلك ، إلا أنه في كثير من الحالات قد يجد المعلومات التي يريدها متوفرة في النشرات الاحصائية التي تصدرها الهيئات الاحصائية العامة في الدولة أو لدى المؤسسات والهيئات والمصالح التي يمكن أن تقوم هي نفسها بجمعها بحكم مقتضيات عملها أو في التقارير التي صدرت في الماضي عن أبحاث شبيهة . إلا أن الباحث في مثل هذه الحالة قد يجد المعلومات التي يريدها مصنفة ومبوبة تبويباً تحكمت فيه اعتبارات إدارية أو فنية أو محاسبية بحيث لا تخدم غرضه تماماً فيضطر الى إعادة تنظيمها وتنسيقها في الشكل الذي يريده ، ولا بد أن يستعين في هذا العمل بالمنهج الاحصائي بالإضافة الى استعانته به عند إجراء التحليل .

ثالثاً : البحوث الاحصائية التجريبية - ويستخدم هذا النوع من البحوث في ميادين مختلفة كالزراعة والطب والصحة العامة والنواحي التربوية والاجتماعية والاقتصادية .

وفي هذا النوع من البحوث يقوم الباحث بالتحكم في الظروف التي تجمع المعلومات على أساسها ، كما أنه يستطيع ان يقوم بتكرار جمع المعلومات اذا

اقتضى الأمر إلى ذلك . وبذلك يختلف هذا النوع عن النوع السابق حيث انه في البحوث التحليلية لا يكون للباحث أية سيطرة على الظروف التي تجمع فيها المعلومات ، كما انه يضطر الى جمعها وقت وقوعها أو في وقت لاحق . ففي دراسة عن الضياع الذي ينتج عن دوران العمل في الصناعة مثلاً ، اما أن يتفق الباحث مع المؤسسات التي تجمع المعلومات منها على تسجيل هذه المعلومات وقت وقوعها أو يقوم في وقت لاحق بجمعها . أما في البحوث التجريبية فان الباحث يواجه عدم التأكد بالنسبة لمشكلة ما ويريد أن يتوصل الى قرار يطمئن اليه ، وبذلك يضطر الى اجراء تجربة قد يعيد إجرائها مرارا قبل التوصل الى القرار الذي يريده . فاذا توصل باحث مثلاً ، إلى نوع جديد من التظعيم ضد مرض معين ويواجه مشكلة الاجابة على السؤال — هل هناك فرق جوهري بين التظعيم بهذا النوع الجديد والتظعيم بالنوع القديم ؟ — للإجابة على هذا السؤال يضطر الى اجراء تجربة على عينتين من البشر تظعم الأولى بالنوع الجديد وتظعم الثانية بالنوع القديم . وحتى يكون الفرق في النتائج راجع فقط الى نوع التظعيم المستخدم لا بد ان تتشابه العينتان في كافة الوجوه ولا بد أن يخضعهما لنفس التأثيرات ، وهو مبدأ أساسي في إجراء التجارب وهو ان تكون وحدات التجربة متساوية تماماً في كل العوامل التي يمكن ان تؤثر عليها ولا تختلف الا في العامل الذي يراد معرفة تأثيره حتى يمكن استبعاد أثر العوامل الاخرى غير هذا العامل .

وبذلك نلاحظ ان الباحث في تجربته يكون في حاجة الى منهاج يسير عليه في تنظيم وتصميم التجربة حتى تكون النتائج علمية صحيحة . وقد اهتم علماء الاحصاء ببحث هذا التنظيم والتصميم ووضعوا القواعد والاصول بحيث أصبحت تكون سوية فرعاً خاصاً في علم الاحصاء يسمى « تصميم التجارب » . لا شك انه في هذا النوع من البحوث يجري العمل بالتعاون من الباحث المختص بالمشكلة نفسها وبين الاحصائي وحتى يتحقق التعاون والتفاهم بينهما لا بد أن يكون كل منهما ملماً ببعض الشيء باختصاص الآخر .

ما هو المنهاج الاحصائي :

تبيننا ما تقدم أن الاحصاء علم قائم بذاته له قواعده وقوانينه الخاصة به ولكن أهميته تظهر في استخدامه كمنهاج للبحث في الميادين العلمية المختلفة . والسؤال الآن ما هي المراحل المختلفة التي يجب أن يمر بها العمل إذا أردنا استخدام هذا المنهاج ؟ ان المنهاج الاحصائي هو ، كما قدمنا ، منهاج علمي وليس منهاج فلسفي . ولذلك تكون أولى مراحلها هي مرحلة جمع المعلومات التي تمثل واقع الظاهرة او الظواهر موضوع البحث حتى تكون المقاييس التي يمكن ان نتوصل اليها فيما بعد بالتحليل تابعة من هذا الواقع وليست مجرد تعبير عن رأي الباحث .

ولا شك ان اول مشكلة تصادف الباحث عند جمع المعلومات هي تحديد المصدر الذي يمكن أن يلجأ اليه . وهناك نوعان من المصادر - المصادر التاريخية ومصادر الميدان - وتشمل المصادر التاريخية الوثائق والمطبوعات الحالية والقديمة . ومن أهم هذه المصادر النشرات الاحصائية التي تصدرها الحكومات والهيئات الخاصة في الدول المختلفة . وتختلف هذه النشرات من دولة الى اخرى فهي قد تزيد بحيث يمكننا أن نجد فيها كل ما نريده من بيانات رقمية كما هو الحال في الدول المتقدمة احصائياً ، وقد تقل فلا نستطيع أن نجد الا اليسير من البيانات التي نريدها كما هو الحال في الدول المتخلفة من ناحية الوعي الاحصائي . ومن المصادر التاريخية النشرات الاحصائية التي تنشرها المنظمات المختلفة التابعة لهيئة الامم والنشرات التي يصدرها مكتب العمل الدولي فيما يختص بإحصاءات العمل .

ويمكن تقسيم المصادر التاريخية الى نوعين : المصادر الاولى وهي التي تقوم بنشرها نفس الهيئة التي قامت بجمع المعلومات وتبويبها مثل جميع

النشرات التي تصدرها الادارات الاحصائية في الدول المختلفة ، والمصادر الثانوية وهي التي تقوم بنشرها هيئة غير التي قامت بجمع المعلومات وتبويبها ، كأن يقوم شخص باقتباس بعض البيانات الموجودة في إحدى نشرات الادارة الاحصائية ويضعها كما هي أو بشيء من التحوير لتناسب الحالة التي يقوم بدراستها وذلك مثل الاحصاءات التي تنشرها الجرائد والمجلات والكتب .

وحيثما يلجأ الباحث الى المصادر التاريخية للحصول على المعلومات التي يريدها يجب أن يأخذ في اعتباره ما يأتي :

١ - أن يلم بجميع التعاريف والمصطلحات التي تكمن وراء الأرقام التي تظهر في هذه النشرات ، حيث أن هذه التعاريف والمصطلحات يمكن أن تختلف من بحث الى آخر وبذلك يكون لها تأثيرها على معنوية الأرقام . فإذا كنا نقوم ببحث عن العمال المتعطلين مثلاً ؛ ولجأنا الى نشرة عن احصاءات العمال لا بد أن نتبين أولاً تعريف المتعطل الذي استخدمته الهيئة التي قامت بأعداد هذه الاحصاءات ، هل قامت باحصاء المتعطلين على أساس أن المتعطل هو كل شخص قادر على العمل وراغب فيه ومستعد لقبوله اذا عرض عليه ولا يجد عملاً ، او على أساس أن المتعطل هو الذي سبق له العمل وتعطل . وغالباً تحتوي مقدمة النشرات الاحصائية على تفسير لجميع التعاريف والمصطلحات التي استخدمت في اعداد الاحصاءات التي تظهر فيها ، وإذا لم تكن واردة في المقدمة يمكن ان نبعث عنها أسفل الجداول في الجزء المخصص للملاحظات .

٢ - ان يراجع الاستمارة التي استخدمت في جمع المعلومات للتأكد من صحة تصميمها ودقة صياغة الاسئلة التي جاءت فيها

٣ - إذا كانت الاحصاءات التي تظهرها النشرات مصدرها دراسات

استقصائية بالمعينة على الباحث ان يتثبت من أن العينة التي سحبت هي من النوع الذي يلائم المجتمع موضوع الدراسة ، ومن أن طريقة سحبها سليمة ، ومن جميع القوانين الرياضية التي استخدمت في تحديد حجم العينة والتقديرات منها وحساب اخطاء هذه التقديرات ، وبشكل عام يجب أن يتثبت من أن الدراسة اجريت وفقاً للقواعد الاحصائية السليمة المتعلقة بموضوع العينات . وعندما تكون الدراسة قد اجريت بالمعينة فان التقرير النهائي عن الدراسة يحتوي غالباً على كل ما يريد الباحث أن يتثبت منه فيما يتعلق بعملية المعينة نفسها .

٤ - إذا كانت الاحصاءات التي تظهرها النشرات مصدرها عمليات ميدان على الباحث أن يقف على المستوى التعليمي للعدادين الذين قاموا بجمع المعلومات وعلى نوع ومدى التدريب الذي حصلوا عليه قبل قيامهم بالعمل وعلى التعليمات التي اعطيت لهم وعلى تنظيم العمل وخاصة بالنسبة لعمليات الاشراف على العدادين ، ومراجعة البيانات ، وجميع العمليات الاخرى التي تجري في المكتب بعد جمع المعلومات من ترميز وجدولة وطبع .

٥ - التأكد من الفترة الزمنية التي تتعلق بها الاحصاءات .

٦ - قراءة عناوين الجداول قراءة فاحصة للتأكد من أن الأرقام الموجودة فيها تتعلق بالدولة بأكملها أو باجزاء منها فقط ، ومن انها خاصة بالفترة الزمنية التي يظهرها عنوان النشرة او بفترة زمنية اخرى ، ومن الاسس التي صنفتم تبعاً لها الأرقام ومن وحدات الأرقام ، ومن انها أرقام كاملة أو مدورة وإذا كانت مدورة فكم رقم محذوف منها (مثلاً ارقام بالمليون أو بالآلف ... الخ) .

٧ - قراءة جميع الملاحظات والتفسيرات الموجودة في أسفل الجداول التي يريدها الباحث .

٨ - قراءة عناوين الأعمدة والأسطر في كعب الجدول وفهمها فهماً جيداً.

٩ - أي رقم وأي عنوان في الجدول يظهر بجانبه إشارة معينة يجب الرجوع الى هذه الإشارة في أسفل الجدول وقراءة الملاحظات المكتوبة أمامها مرة ثانية للامام بما هو وارد فيها .

١٠ - يجب دائماً مراجعة المجاميع والمتوسطات للتأكد منها فلا يجب أن يأخذها الباحث كما هي قبل التحقق منها . ويحسن في بعض الأحيان حساب بعض المتوسطات والمعدلات فهي تساعد في تقييم الأرقام الوارد في الجدول من ناحية الدقة والشمول ، مثلاً يساعدنا حساب المعدلات الحويمة (معدلات المواليد والوفيات ...) في الكشف عن مدى شمول ارقام المواليد والوفيات في مناطق الدولة المختلفة ، كذلك يساعدنا حساب متوسطات أسعار السلع المختلفة في التحقق من ارقام القيم والكميات الواردة في جداول احصاءات التجارة الخارجية .

١١ - يجب قراءة الجدول بعوي وادراك شامل لمحتوياته حتى يستطيع أن تنبئه الى أي شيء يبدو غريباً في الجدول فنحاول ادراك هذه الغرابة بالرجوع الى التفسيرات والتوضيحات في مقدمة النشرة او في أسفل الجدول . أو في التصحيحات الموجودة في نهاية النشرة .

وبذلك يتبين لنا انه من الأفضل الالتجاء الى المصادر الأولية حيث انها تحتوي في الغالب على التفسيرات والتوضيحات المطلوبة وعلى جمع المستندات الخاصة بعملية جمع المعلومات مثل الاستمارة التي استخدمت والتعليمات التي اعطيت للعدادين وكذلك على وصف شامل لجميع خطوات العمل ولا يعني ذلك ان المصادر الثانوية لا تحتوي على مثل هذه الأشياء ، ففي بعض الأحيان تنقلها هذه المصادر من مصادرها الأولية ، الا انه غالباً ما يؤدي النقل الى اخطاء في الارقام او في التفسيرات . ومن جهة اخرى قد يكون من الأفضل الالتجاء الى المصادر الثانوية اذا كانت هذه المصادر اعادت نشر الاحصاءات في

شكل يكون اكثر فائدة للباحث في بحثه وفي هذه الحالة تتوقف اهمية المعلومات التي تظهر فيها على مدى كفاية القائمين بهذا العمل وخبرتهم في موضوع البحث .

وعندما لا تتوفر المصادر التاريخية او عندما لا تكون كافية لاجراء البحث يكون الباحث مضطراً الى الالتجاء الى الميدان للحصول على المعلومات التي يريدها ، وفي هذه الحالة يسمى البحث ببحث ميدان - ويقوم هذا النوع من البحوث اساساً على تصميم استمارة تحتوي على اسئلة توجه الى المستجوبين المعنيين للحصول على اجاباتهم التي تكون هي بمثابة المعلومات التي يريدها الباحث . ويواجه الباحث في هذا النوع من البحوث مشاكل وصعوبات جمة ، الأمر الذي يجعلها في حاجة الى تنظيم محكم دقيق لجميع مراحل العمل الميداني والمكتبي وذلك لضمان دقة المعلومات التي يمكن التوصل اليها وضمان انجاز العمل في الوقت المناسب ، واذا تذكرنا ان جمع المعلومات بهذه الطريقة يؤدي الى الاحتكاك بالجمهور او المنشآت والهيئات المختلفة واذا تذكرنا ان الوعي الاحصائي غالباً ما يكون ضعيفاً بل انه يمكن ان يكون سلبياً في بعض الأحوال حيث يواجه الباحث رفض غالبية المستجوبين الاجابة أو تعمد عدم اعطاء الاجابات الخاطئة المضللة ، اذا تذكرنا ذلك ادر كنا مدى ما يحتاجه الباحث من تنظيم للتغلب على هذه الصعوبة بالدرجة الأولى وعلى الصعوبات الاخرى التي يمكن أن تنشأ عن أوجه .النقص في جهاز الموظفين القائم بالعمل . وفي الفصل التالي نناقش بشيء من التفصيل خطوات العمل المختلفة التي يتم وفقاً لها جمع المعلومات من الميدان .

بعد أن تم مرحلة جمع المعلومات تبدأ المرحلة الثانية في المنهاج الاحصائي وهي الخاصة بعرض هذه المعلومات عرضاً يساعد قياً بعد على تحليلها . ويكون عرض المعلومات بواسطة تصنيفها وتبويبها في جداول ثم بتوضيحها بالرسومات البيانية كلما كان ذلك ممكناً . ولا شك أن عملية التصنيف والتبويب لا

تأخذ مجهوداً كبيراً من قبل الباحث الذي يكون مصدر بحثه المصادر التاريخية المختلفة . ولا يجب أن يتبادر الى ذهننا ان الباحث في مثل هذا النوع من البحوث يأخذ المعلومات كما هي مبوبة في النشرات ، حيث انه في الغالب يضطر الى إعادة تنسيقها وتبويبها لتنسيق والتبويب الذي يتفق والبحث الذي يقوم به . اما في بحوث الميدان فان المعلومات تأتي الى المكتب في استمارات يصل عددها في بعض التعدادات (تعداد السكان مثلاً) الى الملايين والمعلومات كما تظهر في الاستمارات لا يمكن بأي حال من الأحوال أن تساعد في فهم اتجاه الظاهرة أو الظواهر موضوع للبحث ، بل لا يمكن أن يستوعبها الباحث في ذهنه مهما أوتي من الذكاء وبذلك تكون عملية التصنيف والتبويب عملية واجبة حتى يمكن أن يأخذ المنهاج الاحصائي مجراه وحتى يمكن أن ينتهي بنا الى النتائج التي نقوم بالبحث من أجلها ، ذلك لاننا وان كنا نجمع المعلومات عن كل وحدة على حدة ، الا ان هذه المعلومات الفردية لا تهتمنا في شيء حيث ان اهتمامنا يتجه الى معرفة الاتجاهات العامة للظواهر موضوع البحث . فاذا كنا نقوم بدراسة عن مستوى الاجور ، مثلاً ، فانتا قد نسأل كل عامل على حدة عن أجره ، وسؤالنا هذا لا يعني اننا نرغب في معرفة أجر كل عامل فعلاً وانما لاننا نستطيع أن نتوصل عن طريق معرفة اجور العمال الى تحديد المتوسط العام لأجرهم وتوزيعهم في فئات معينة للأجر وغير ذلك من الحقائق التي نرغب في معرفتها عن ظاهرة الأجر بشكل عام . ولهذا عند تبويب الأجور في فئات معينة فان الاجور تفقد صفتها الشخصية وتصبح في نظراً مجرد ظاهرة تجري عليها تحليلنا الذي نريده .

وبعد أن تتم مرحلة عرض المعلومات تنتهي بذلك البحوث الاحصائية الوصفية حيث لا يتبقى الا طبع النشرة التي تحتوي على الجداول والرسوم التي أمكن التوصل اليها . اما في البحوث التحليلية والتجريبية حيث يكون الهدف الأساسي هو التوصل الى المقاييس المختلفة التي تدل على اتجاهات الظواهر

موضوع البحث تكون المرحلة التي تتلو عرض المعلومات هي اجراء عمليات التحليل المختلفة التي يرغب الباحث في اجرائها . والتحليل الاحصائي يشمل عمليات كثيرة ندرس بعضاً منها في هذا الكتاب مثل حساب المتوسطات والتشتت والالتواء والارتباط والانحدار ... الخ

ويضيف بعض الكتاب إلى المراحل الثلاث السابقة مرحلة رابعة هي التفسير أي توضيح مدلولات المقاييس المختلفة التي أمكن التوصل إليها بالتحليل، والواقع ان التفسير وان كان ينبغي على المقاييس التي يمكن التوصل إليها باستخدام المنهاج الاحصائي إلا انها عملية ليست ذات طبيعة احصائية بحتة حيث انها تحتاج الى خبرات لا تكون خبرات احصائية مجردة ، إذ ان تحديد مدلولات المقاييس يتوقف على المعرفة المتخصصة بالناحية العلمية التي يعمل فيها الباحث . فلو توصلنا مثلاً ، باستخدام المنهاج الاحصائي الى مقياس معين للارتباط بين لون الشعر ولون العينين فان ادراك هذا الارتباط وتفسيره ليس من واجب الاحصائي .

التضليل الاحصائي :

تبين لنا ان استخدام المنهاج الاحصائي يوصلنا الى نتائج رقمية ذات قوة اقناع لا حد لها حيث ان الارقام أصبحت تمثل في ذهن انسان العصر الحديث الحقيقة التي لا يمكن مناقشتها ، إلا اننا نلاحظ أن تفسير النتائج الرقمية التي نصل إليها عن طريق استخدام المنهاج الاحصائي يمكن أن يكون مضللاً ما لم نعالج هذا التفسير بحكمة ومهارة وما لم نكن موضوعيين في تعليقاتنا على هذه النتائج ، وليس ادل على ذلك من النقاش الذي يحدث بين فريقين من الناس فيما يتعلق بموضوع معين وكل منهما يؤيد رأيه ببيانات احصائية قد تكون في بعض الاحيان مستقاة من نفس النشرة . والسؤال الآن ما هي الاسباب التي يمكن أن تؤدي الى ذلك ؟ هناك أسباب كثيرة نورد بعضها في الآتي :

١ - الجهل بالعاريف والمصطلحات التي تكمن وراء الأرقام . فعند مقارنة البطالة في دولة ما في فترتين زمنيتين ، مثلاً ، يمكن أن تقع في خطأ فادح فنقول ان البطالة في هذه الدولة قد ازدادت حيث ان الأرقام تدل على ذلك مع اننا لو حاولنا ان ندرك ما وراء هذه الأرقام وأدركناه ادراكاً صحيحاً لتوصلنا الى عكس النتيجة السابقة بالرغم من الدلالة الظاهرية للأرقام ، اذ قد تكون الأرقام في الفترة الأولى مبنية على تعريف للمتعطّل يقتصر على اعتبار الشخص متعطلاً اذا كان قد سبق له العمل وتعطل ، كما انها قد تكون مأخوذة من مكاتب التشغيل ونحن نعلم علم اليقين أن القانون لا يجبر كل متعطّل على تسجيل اسمه في هذه المكاتب وبذلك يوجد عدد كبير من المتعطّلين الذين لا يسجلوا اسمائهم اعتماداً على قدرتهم على أن يجدوا عملاً لأنفسهم دون الالتجاء الى مكاتب التشغيل ، بينما في الفترة الثانية يمكن أن تكون الأرقام مبنية على تعريف أكثر شمولاً حيث اعتبر الشخص متعطلاً سواء سبق له العمل او لم يسبق له ذلك طالما انه قادر على العمل ويبحث عنه دون أن يجده ، كما انها قد تكون نتيجة لدراسة استقصائية شاملة لكل فئات المجتمع ولذلك كله تبدو الأرقام في هذه الفترة اكبر منها في الفترة الاولى .

كذلك يمكننا ان نناقش فنقول ان مستوى الاجور في دولة ما قد ارتفع خلال فترة معينة بينما يناقش فريق آخر من الناس فيقول ان العكس هو الصحيح ، وكل منا يؤيد رأيه ببيانات احصائية نستقيها جميعاً من النشرة الاحصائية الرسمية . واذا اردنا ان ندرس الموضوع دراسة موضوعية امكنا ان نتبين ان رأي كل من الفريقين صحيح اذا اخذنا في الاعتبار نوع البيانات الاحصائية التي تؤيد الرأيين - فقد يكون الفريق الأول معتمداً في رأيه على بيانات خاصة بكسب العمل ويمكن ان تكون قد ارتفعت بسبب زيادة ساعات العمل الاضافي او بسبب زيادة الاجور العينية التي تمنحها المؤسسات للمستخدمين ، بينما يكون الفريق الثاني معتمداً في رأيه على بيانات خاصة

بمعدلات الاجور التي تكون قد انخفضت بسبب استخدام الآلات الحديثة التي تجعل العمل أكثر آلية واقل عناء .

٢ - عدم معرفة مزايا وعيوب الطرق المختلفة التي يمكن ان تستخدم في جمع المعلومات التي سوف يجري تحويلها الى احصاءات بالتصنيف والتبويب . فاذا كنا ندرس الحالة الصحية في دولة ما خلال فترة طويلة من الزمن فقد تفاجئنا الارقام بازدياد عدد الامراض التي تظهرها الجداول ، الأمر الذي يجعلنا نصل الى استنتاج خاطيء فنقول ان الحالة الصحية في هذه الدولة في تدهور مستمر . ولكننا اذا ادركنا ان مصدر الأرقام الخاصة بالامراض هو المستشفيات والمؤسسات الصحية المختلفة استطعنا أن نفهم سبب زيادة هذه الارقام حيث ان توسع الدولة في انشاء المستشفيات والمؤسسات الصحية في المناطق المختلفة يؤدي الى التبليغ عن امراض لم يكن من الممكن التبليغ عنها عندما كانت هذه المناطق محرومة من جميع الخدمات الصحية

كذلك يمكن أن يكون مصدر المعلومات الدراسة الاستقصائية بالمعينة فاذا لم ندرك ان التحيز يمكن أن يتطرق الى مثل هذه الدراسات اذا لم تتبع الاصول الاحصائية السليمة ، فاننا لا بد وأن نقع في أخطاء فادحة اذا اعتمدنا على نتائج هذه الدراسات دون أن نتأكد من اتباع هذه الاصول . وفي بعض الاحيان قد يكون التحيز متعمداً بقصد التضليل ، والتحيز المتعمد يمكن أن يكون واعياً ويمكن أن يكون لا شعورياً . ففي كثير من الدراسات التي يكون هدفها الدعاية يكون التحيز متعمداً والمثل على ذلك ما يقوم به بعض الجرائد والمؤسسات التليفزيونية من دراسات استقصائية لا ثبات اراء يؤيدونها سلفاً . الا انه في بعض الأحيان يكون الباحث متأثراً بإحساس معين قائم باطنياً ، الأمر الذي يجعله في تصميم استمارة البحث يصيغ الاسئلة بشكل يعطيه اجابات معينة يؤيدها هو نفسه ، والمثل على ذلك دراسة استقصائية عن أسباب فشل الزواج يقوم بها رجل فشل في زواجه ويلقى لذلك كل التبعة على المرأة ، فانه لا شعورياً يعمل على إثبات هذه النتيجة بأية وسيلة .

٤ - عدم جمع المعلومات الكافية عن الظاهرة أو الظواهر موضوع البحث ، ففي دراسة عن الارتباط بين درجات الطلبة في مادة الاحصاء ودرجاتهم في مادة المحاسبة يمكن ان نصل الى نتيجة مضللة اذا اعتمدنا في دراستنا على عينة صغيرة من عشرة طلبة مثلاً حيث أنه في هذه الحالة يكون احتمال تأثير الصدفة على النتيجة احتمال كبير جداً . وفي بحث عن أفضل دواء جديد في معالجة مرض معين لا يمكن أن نشق بالنتيجة اذا كانت مبنية على تجربة العقار على عدد قليل جداً من الأشخاص . وخلاصة القول إن نتائج الدراسات بالمعينة تعتمد الى حد كبير على حجم العينة أي على عدد الوحدات التي تتكون منها العينة وسوف نتبين عند دراسة موضوع العينات ان هناك علاقة عكسية بين حجم العينة ومقدار الخطأ المحتمل حدوثه في النتائج . ولا يجب أن يتبادر الى الذهن ان مجرد تكبير حجم العينة فيه الضمان الكافي لدقة نتائج العينة ، فهناك قواعد وأصول كثيرة خاصة بهذا الموضوع التي لا بد من أتباعها لضمان دقة تمثيل العينة للمجتمع الذي سحبت منه .

٥ - عدم معرفة مضمون الفئات المختلفة التي صنفت تبعاً لها البيانات الاحصائية ويحدث هذا غالباً عند مقارنة أرقام التجارة الخارجية لدولتين أو لدولة واحدة في فترتين مختلفتين ، ففهوم كلمة مواد أولية أي مكوناتها يمكن أن تختلف من دولة الى أخرى أو من وقت الى آخر في نفس الدولة ، ولعل ذلك هو السبب في قيام المكتب الإحصائي التابع لهيئة الأمم بوضع تصنيف موحد للبضائع التي يجري تبادلها فيما بين الدول المختلفة حتى تكون المقارنة بين احصاءاتها مقارنة صحيحة غير مضللة . كذلك يحدث هذا الخطأ عند مقارنة التوزيع النسبي للسكان تبعاً للمهن أو تبعاً للنشاط الاقتصادي لاختلاف مفهوم الفئات المختلفة التي تدخل في تصنيف المهن أو في تصنيف النشاط الاقتصادي ، ولعل ذلك هو السبب كذلك في ان الهيئات الدولية المتخصصة (مكتب العمل الدولي ومكتب الإحصاء التابع لهيئة الأمم) قد وضعت تصنيفات موحدة وأوصت باستخدامها . كذلك يحدث هذا الخطأ عند مقارنة احصاءات

الأمراض المختلفة حيث يمكن ان يختلف مفهوم الأمراض الصدرية أو الأمراض الباطنية من دولة الى اخرى أو من وقت الى آخر ولذلك وضعت منظمة الصحة العالمية تصنيفاً موحداً للأمراض والايذاءات وأسباب الوفاة وأوصت باستخدامه في الدول الأعضاء حتى يمكن مقارنة احصاءاتها مقارنة علمية صحيحة . وقد بدأ كثير من الدول فعلاً في استخدام هذه التصنيفات الدولية.

٦ - كذلك يمكن أن نقع في أخطاء فادحة بسبب الخطأ في تفسير النتائج التي نصل اليها . ويحدث ذلك أما بسبب اهمال أحد العوامل التي يمكن أن يكون لها تأثير على الظاهرة والظواهر موضوع البحث أو بسبب عدم الالمام بموضوع البحث المأما كافياً . فاذا كنا نقوم بتجربة لاختبار أثر عامل معين لا بد وأن نجعل باقي العوامل الاخرى التي يمكن أن يكون لا تأثير على هذه الظاهرة مشتركة كما سبق أن قدمنا ، فاذا أهملنا أحد العوامل الاخرى وأبقيناه مختلفاً خلال اجراء التجربة فان النتيجة لا يمكن أن ترتبط بالعامل الذي ندرسه فقط . كذلك اذا كنا ندرس الارتباط بين الاتجاه العام لسلسلتين زمنيتين ، مثلاً الارتباط بين الاتجاه العام لصادرات الدولة في عدة سنوات والرقم القياسي لاسعار الجملة في هذه السنوات ، ولم نخلص السلسلتين من أثر الانواع الاخرى من التغيرات التي يمكن أن تؤثر عليها (التغيرات الموسمية والدورية والعرضية) فان النتيجة التي نصل اليها لا يمكن أن تكون ذات دلالة صحيحة على وجود ارتباط بين الاتجاه العام للظاهرتين . كذلك اذا كنا نقارن مستوى الاجور في صناعتين مختلفتين وقبين من النتائج الرقمية أن المستوى في أحدهما أعلا من الآخر ، فإتنا يمكن أن نقع في خطأ فادح إذا فسرنا ذلك نتيجة اختلاف الصناعتين ، حيث يمكن أن يكون السبب في الاختلاف ليس راجعاً الى اختلاف نوع الصناعة بقدر ما هو راجع إلى أن احدي الصناعتين تتركز في المدن حيث مستوى الاجور أعلى والاخرى تتركز في الريف حيث مستوى الاجور أقل . كذلك إذا كنا في صدد مقارنة مستوى المعيشة في دولة ما في فترتين زمنيتين واعتمدنا في ذلك على أرقام

الدخل الشهري فانتنا يمكن ان نفع في اخطاء فادحة اذا لم نأخذ في اعتبارنا نمو السكان خلال الفترة موضوع البحث ، وتطور الاسعار خلال هذه الفترة ، واختلاف توزيع الدخل بين السكان حيث يمكن ان يكون التوزيع في الفترة الاولى اقل عدالة منه في الفترة الثانية .

تبين لنا من هذه المناقشة ان هناك أسباب كثيرة يمكن أن تجعل أحكامنا على الظواهر التي نبحثها احكاما خاطئة ، ويحدث ذلك غالباً عندما لا يكون الانسان ملماً بالمنهاج الاحصائي المأمأ كافياً . ولذلك فدراستنا لهذا المنهاج سوف تساعدنا بالاضافة الى معرفة كيفية استخدامه في البحث العلمي ، في ألا نفع في الاخطاء الفادحة التي يمكن ان نفع فيها بسبب جهلنا به خاصة إذا كنا نقوم بابحاث في الميادين العلمية التي تهتمنا .

الفصل الثاني

جمع المعلومات

في هذا الفصل ناقش الخطوات المختلفة التي يجب علينا أن نتبعها إذا أردنا النزول الى الميدان لجمع المعلومات التي تلزمننا في البحث الذي نريد القيام به . وقبل أن نبدأ هذه المناقشة يجب أن اشير الى ان هناك كثير من الادارات الحكومية التي يتجمع لديها معلومات كثيرة في السجلات التي تمسكها فيها يختص بالاعمال التي تقوم بها ، وهي تقوم بذلك ليس لغرض احصائي وانما لأغراض اخرى يمكن أن تكون اغراض ادارية أو أغراض قانونية أو أغراض مشتركة بينها ؟ فمديرية الجمارك مثلا تسجل في دفاتها معلومات مختلفة عن كل بضاعة تخرج من الدولة أو تدخل اليها ، ومديرية الاحوال الشخصية تقوم كذلك بتسجيل معلومات مختلفة عن كل مولود وكل وفاة وكل زواج وكل طلاق ، وأقسام البوليس تقوم ايضا بتسجيل معلومات مختلفة عن كل حادث طريق يحدث في الدولة . الخ .. هذه المعلومات كما هي مسجلة في سجلات هذه المديريات المختلفة لا تعتبر احصاءات ، ذلك لانها معلومات فردية ، أي عن كل حالة على حدة ولكنها تعتبر مصدرا هاما للإحصاءات الخاصة بهذه الموضوعات المختلفة وذلك يكون بتصنيفها وتبويبها في جداول تظهر ارقاما مجمعة . والمهم في هذا الصدد ان المعلومات التي تسجلها هذه المديريات المختلفة قد تكون مسجلة بشكل لا يساعد على تحويلها الى احصاءات بالتصنيف والتبويب ، أو قد تكون غير كافية لاعداد احصاءات عن هذه

الموضوعات تساعد الباحث في بحث معين يقوم به . فاذا كنا نقوم ببحث عن خصوبة المرأة في فئات العمر المختلفة مثلاً ، نكون في حاجة الى بيان عن عمر الام لكل مولود فاذا لم تكن مديرية الأحوال الشخصية تهتم بتسجيل هذا البيان فانتا لا نستطيع أن نستمر في بحثنا حيث لن نجد المعلومات التي نريدها . وبذلك يكون من الأفضل لو تعاونت هذه المديريات المختلفة مع مديرية الاحصاء او مع الأقسام الاحصائية في الوزارات المختلفة على تصميم الاستمارات التي تجمع بواسطتها المعلومات الخاصة بهذه الموضوعات وبذلك يمكن تضمينها للأسئلة المختلفة التي تهتم الباحثين في اجابتهم ، كما يمكن تصميم هذه الاسئلة بالشكل الذي يساعد في تطبيق المنهج الاحصائي .

كذلك نلاحظ ان خطوات العمل عندما يكون الميدان هو مصدر المعلومات التي نريد جمعها لا تختلف سواء أردنا ان نجتمع المعلومات بالعد الشامل أي من جميع وحدات المجتمع الاحصائي ^(١) موضوع البحث او بالمعينة أي من بعض وحدات هذا المجتمع الذي سحبت منه بحيث تكون مع بعضها صورة صادقة له ، والاختلاف الوحيد بالاضافة الى الخطوات الخاصة بالمعينة والتي تقوم بها فقط عندما تكون الدراسة بهذه الطريقة ، هو في نطاق العمل حيث يكون النطاق واسع جداً في التعدادات (العد الشامل) وأضيق بكثير

(١) يقصد بالمجتمع الاحصائي مجموعة الوحدات موضوع البحث ويجب ان تكون هذه الوحدات معروفة بصورة واضحة بحيث يمكن تمييزها عن غيرها من الوحدات التي تكون مجتمعة آخر . فاذا كنا بصدد بحث عن الموظفين المصنفين في لبنان ، يكون كل موظف مصنف في الحكومة وحدة احصائية ومجموع الموظفين المصنفين يكونوا المجتمع الاحصائي . واذا كنا بصدد دراسة عن الصناعة في لبنان فان مجموع المؤسسات الصناعية تكون هي المجتمع الاحصائي وكل مؤسسة منها هي الوحدة الاحصائية . أما اذا كنا بصدد دراسة عن صناعة الغزل والنسيج فقط فان مجموع المؤسسات التي تعمل في هذا النشاط تكون هي المجتمع الاحصائي وكل مؤسسة منها هي الوحدة الاحصائية .

والجتمعات الاحصائية تكون دائمة التغير من وقت الى آخر وهذا ما يجعل الدراسة الاحصائية لأي مجتمع مرتبطة بالوقت الذي جمعت فيه المعلومات الخاصة بهذه الدراسة .

في العينات . وقد يبدو هذا الاختلاف بسيطاً ، إلا أن ما يترتب على ذلك من ناحية الدقة في النتائج امر جوهري . ذلك لأنه عندما يكون العمل على أساس العد الشامل تزداد نسبة الخطأ بسبب صعوبة التنظيم والاشراف الدقيق ، بينما اذا كان العد (جمع المعلومات) بالمعاينة فان الجهاز القائم بالعمل يستطيع ان يتحكم في التنظيم والاشراف بحيث يمكن أن ينقص الاخطاء الى درجة كبيرة . ولعل ذلك هو السبب في ان كثيراً من الاحصائيين يعتقدون ان الدراسة بالمعاينة وخاصة بالنسبة للموضوعات المعقدة تأتي بنتائج أكثر دقة من الدراسة بالعد الشامل .

وتقودنا هذه المناقشة الى بحث موضوع الاخطاء التي يمكن أن تظهر في نتائج الابحاث الاحصائية والتي يجب أن يعمل الجهاز القائم بالعمل على تلافيها أو انقاصها الى ادنى حد ممكن ، ولن يتأتى له ذلك إلا بالتنظيم المحكم والاشراف الدقيق والمراجعة المستمرة لنتائج كل خطوة من خطوات العمل ، وبما لا شك فيه ان الدراسة بالمعاينة تهيأ لنا فرصة أفضل لتحقيق كل ذلك بسهولة ويسر .

ان أول نوع من الاخطاء التي يمكن ان تظهر في الابحاث الاحصائية هي الاخطاء في اجابات المستجوبين . ويمكن ان تكون هذه الاخطاء متممة منهم ، الامر الذي يجعل البحث يفقد قيمته فقداناً تاماً وبذلك يتحتم على الباحث اما أن يحاول تصحيح هذه الاخطاء بطريقة ما ويمكن أن يتم ذلك بأخذ عينة من المستجوبين وارسال عدادين مدربين تدريباً عالياً للاتصال بهم وتنويع الاسئلة في الاستمارة بحيث يمكن الكشف عن أخطاء المستجوب عند اجابته على سؤال معين ومراجعته بعد تنبيهه الى ذلك . ويحدث هذا النوع من الخطأ عندما لا تكون هنالك الدعاية الكافية للبحث قبل النزول الى الميدان لجمع المعلومات وخاصة في المجتمعات التي ينعدم فيها الوعي الاحصائي ، الامر الذي يجعل المستجوبين يسيئون الظن بالأسئلة التي توجه اليهم فيتمعدون

الاجابة الخاطئة (السؤال عن دخل الأسرة مثلاً) ، كما ان الخوف من الحسد في بعض الدراسات يدفع المستجوبين أيضاً الى الاجابات الخاطئة (السؤال عن عدد أطفال الأسرة أو عدد الماشية والدواجن التي تملكها الأسرة) . وفي كلتا الحالتين تؤدي الدعاية للبحث خاصة اذا اتبعت جميع الوسائل الممكنة الى اتخاذ المستجوبين موقفاً لا يشوبه أي خوف من الاجابة الصحيحة .

كذلك يمكن أن تظهر الأخطاء في اجابات المستجوبين دون أي تعمد منهم وتكون بذلك أخطاء عشوائية سواء بالزيادة أو النقص ، الأمر الذي لا يغير كثيراً من المتوسطات التي نرغب في التوصل اليها فيما بعد باجراء التحليل الاحصائي على المعلومات التي نجمعها . على أن ذلك لا يجب أن يجعلنا نتهاون في جميع المعلومات الصحيحة حيث اننا لا نستطيع في حالة معينة ان نقرر ما اذا كانت أخطاء المستجوب أخطاء متعمدة أو عشوائية وحتى اذا اعتقدناها سوف تكون عشوائية لا نستطيع أن نقرر مداها . لذلك يتحتم علينا انقاص هذه الأخطاء إلى أدنى حد ممكن ، وحتى نستطيع أن ندرك الوسائل التي يمكن اتباعها في هذا الشأن يجب أن نتعرف على الأسباب التي يمكن أن تؤدي الى هذا النوع من الأخطاء . هناك أسباب كثيرة يمكن أن تؤدي الى ذلك ، منها سوء تصميم استمارة البحث بحيث تأتي الاسئلة في صيغ لا يفهمها المستجوب فهماً كاملاً صحيحاً ، كما ان لشخصية العدادين الذين يقومون بالاتصال بالمستجوبين ولتدريبهم على العمل اثر كبير على الاجابات التي يمكن التوصل اليها . كذلك يمكن أن يكون عدم شمول الحصر لوحدات المجتمع موضوع البحث قبل النزول الى الميدان لجمع المعلومات سبباً في أن تكون الاجابات التي تحصل عليها غير شاملة لجميع وحدات المجتمع ، الأمر الذي يؤدي الى تسرب الخطأ الى النتائج خاصة اذا كانت الوحدات التي لم يشملها الحصر ذات صفة مميزة لها علاقة بموضوع البحث . لذلك يكون للتصميم الجيد للاستمارة ، ولانتقاء العدادين الاكفاء وقدرتهم التدريب الكافي

وللاشراف عليهم اثناء قيامهم بالعمل لمساعدتهم في مواجهة المشاكل التي
يحدث ان تصادفهم في هذا الوقت ، ولحصر وحدات المجتمع حصراً شاملاً
او كبير على كمية ونوعية الاجابات التي نحصل عليها .

ومن ناحية اخرى يحتمل ان يجيب المستجوب اجابة صحيحة ولكن
العداد نفسه يسجلها خطأ في استمارة البحث ، ويحدث ذلك اما بسبب سرعة
العمل او عدم تدريب العدادين التدريب الكافي على كيفية كتابة الاجابات
في الاستمارة أو عدم الاشراف الدقيق عليهم الأمر الذي يحتمل لا يهتمون
بالاتصال بالمستجوبين للحصول على اجاباتهم وانما يكتبون هم انفسهم الاجابات
التي يتخلونها رغبة منهم في عدم تحمل مشاق العمل في الميدان. كذلك يمكن
أن يكون السبب في هذا النوع من الأخطاء عدم تصميم الاستمارة التصميم
بحيث لا يكون فيها الفراغات الكافية لكتابة الاجابات ، الامر الذي يحتمل
العدادين يعملون على اختصار الاجابات المكتوبة اختصاراً مخلاً فلا يكون
لما هو مكتوب في الاستمارة اية قيمة عند اجراء عمليات التبويب ، ولا شك
ان تدريب العدادين وشعورهم بالمسئولية ودفع الاجور المحزية لهم وتأكدتهم
من مراجعة المشرفين للاستمارات التي يذهبون منها وعدم قبضهم أي أجر عن
الاستمارات غير الصحيحة ، ومعاقبتهم بالطرد من العمل اذا ظهر ان الاجابات
التي كتبوها في الاستمارات مزيفة ، كل هذه الوسائل يمكن بواسطتها انقاص
هذا النوع من الاخطاء الى أدنى حد ممكن .

وعندما يكون البحث بالمعاينة يمكن أن تظهر أخطاء التسجيل بشكل
خطير بحيث يفقد البحث قيمته من ناحية تحيز النتائج التي يمكن التوصل اليها.
ذلك انه في مثل هذه البحوث تسحب بعض وحدات المجتمع بطريقة معينة
بحيث لا يكون هناك أي تحيز في اختيارها حتى يمكن أن تكون مع بعضها
صورة صادقة للمجتمع الذي تمثله . وعندما ينزل العدادون الى الميدان تعطي
لهم عناوين هذه الوحدات ويطلب منهم الاتصال بها لجمع المعلومات المطلوبة ،

ولكن يحدث أحياناً ان لا يوجد المستجوب في العنوان المحدد لأسباب مختلفة منها تغيير عنوانه أو وجوده خارج مسكنه وقت زيارة العداد ، وكذلك يحدث أحياناً أن يرفض المستجوب الاجابة رفضاً باتاً ، فاذا لم تكن هناك التعليقات المشددة التي تطالب العدادين بعدم التصرف في مثل هذه الأحوال الا بعد استشارة الجهاز القائم بالبحث واذا لم يكن هناك الاشراف والمراجعة الدقيقة فان العدادين جهلاً منهم بما سوف يترتب على تصرفهم من أخطاء فادحة يستعصون عن هذه الوحدات بوحدات اخرى يختارونها هم انفسهم دون أي اهتمام . ولا شك ان هذه الوحدات غالباً ما تكون من نوع معين ، (مثلاً اسر مكونة من رجل أعزب يسكن وحده أو رجل وزوجته والاثان يعملان) الأمر الذي يحمل هذا النوع غير ممثل في العينة وبذلك لا تكون نتائج البحث مصورة تصويراً دقيقاً لخصائص المجتمع موضوع الدراسة . ولا يمكن معالجة هذا النوع من الاخطاء الا بتقنية العدادين عند تدريبهم الى معنى العينة وأهمية تمثيلها للمجتمع الذي تسحب منه وخطورة ما يترتب على تصرفهم السابق من هذه الناحية ، هذا بالإضافة الى التعليقات المشددة التي توجه اليهم ، وحزم الجهاز القائم بالبحث في تطبيق العقوبات المنصوص عليها في هذه التعليقات ، والاشراف على العدادين أثناء عملهم في الميدان واخيراً المراجعة الدقيقة لما يقومون به من عمل .

والنوع الثالث من الأخطاء هي الاخطاء التي يمكن أن تظهر أثناء عمليات التبويب خاصة عند استعمال الآلات الاحصائية . ان استخدام الآلات الاحصائية في التبويب يتطلب عمليات مختلفة مثل الترميز والتثقيب والفرز والتبويب ، وفي كل عملية منها يمكن أن يحدث خطأ معين ولا سبيل الى انقاص هذه الأخطاء الا بمراجعة كل عملية مراجعة دقيقة واستبعاد الموظفين الذين تتكرر الأخطاء فيما يقومون به من عمل وبذلك يقتصر الجهاز على استخدام الموظفين الأكفاء . هذه الأنواع الثلاثة من الأخطاء يمكن أن نسميها بالأخطاء العامة لأنها تظهر في كل دراسة احصائية سواء كانت بالعد الشامل أو بالمعينة ، والصفة

الأساسية لهذه الأخطاء انه لا يمكن القضاء عليها تماماً حيث ان ذلك يعني ان كل مستجوب قد أعطى الاجابة الصحيحة الدقيقة وكل عداد قد قام بعمله على أكمل وجه ، وكل موظف في المكتب يعمل بدقة كاملة . ان البحوث الاحصائية هي من النوع الذي لا يمكن ان يقوم بها شخص واحد وحده فهي عمل مجموعة من الناس ، كما انها في الغالب تتطلب الاحتكاك بالجمهور ، ولذلك لا يمكن أن نتصور قيام مجموعة كبيرة من الناس (خاصة في التعدادات) بالأعمال التي توكل اليهم بدقة كاملة. ولذلك فان القيام بالبحث على نطاق ضيق (بالمعينة) يساعد على انقاص هذه الأخطاء الى أدنى حد ممكن وبذلك تزداد دقة النتائج . واذا كانت هذه الأخطاء لا يمكن القضاء عليها تماماً فان انقاصها الى أدنى حد ممكن أمر يجب أن يكون هو الهدف الأول للجهاز القائم بالبحث ، وكما سبق أن ذكرنا يكون التنظيم الحكم هو السبيل الوحيد الى ذلك.

على انه يجب ان نلاحظ ان اخطاء الاجابة يمكن التخلص منها تماماً اذا أمكن الحصول على المعلومات المطلوبة بطريقة موضوعية ، وفي بعض البحوث يمكن التوصل الى ذلك عن طريق الحصول على المعلومات بالقياس الفعلي . فاذا كنا نقوم ببحث عن أطوال وأوزان الطلبة نستطيع عدم توجيه السؤال اليهم . ونقوم فعلاً بالقياس لهاتين الظاهرتين . كذلك اذا كنا نقوم ببحث عن المساحة المزروعة قمحاً نستطيع عدم توجيه السؤال الى حائزي الأرض عن هذا البيان ونقوم بدلاً من ذلك بقياس المساحة . بهذه الطريقة نتخلص من أخطاء الاجابة الا اذا كان هناك خطأ في المقياس الذي نستخدمه أو اذا كان الشخص القائم بالقياس غير مدرب تدريباً كافياً على استخدام اجهزة القياس التي يستعملها .

خطوات جمع المعلومات :

أشرت الى الاخطاء السابقة قبل أن أناقش خطوات جمع المعلومات حتى

يتبين لنا أهمية تنظيم هذه الخطوات في شكل متناسق متصل ، بل ان كثير من هذه الخطوات هو في الواقع نتيجة حتمية لرغبة الجهاز القائم بالبحث في انقاص الأخطاء العامة الى أدنى حد ممكن

ان أول ما يجب أن نلاحظه ان جمع المعلومات لا يمكن ان يتم على نطاق واسع (التعدادات) بدون أساس قانوني يركز عليه . بل انه في بعض الدول لا يمكن لاية ادارة حكومية أو غير حكومية أن تقوم بعمل احصائي فيه احتكاك بالجمهور إلا إذا كان القانون يسمح لها بذلك . لهذا يوجد في معظم دول العالم في الوقت الحاضر تشريع احصائي ينص على الآتي :

١ - حق مديرية الاحصاء وغيرها من الأجهزة الاحصائية الحكومية في جمع المعلومات . وفي بعض الدول لا ينص القانون على هذا الحق العام وانما لابد أن يصدر مرسوم خاص يحدد المعلومات التي سوف تقوم بجمعها مديرية الاحصاء وبناء على ذلك يحدد المستجوبين الذين تقع عليهم مسؤولية اعطاء المعلومات . ان النص القانوني على هذا الحق يقوى من مركز العداد ازاء المستجوب حيث لا يظهر أمامه مستظلاً لاختباره وأسراره بل يقوم بعمله بناء على حق قانوني. وفي ذلك بعض الحافر للمستجوب على الاجابة على الأسئلة التي توجه اليه .

٢ - واجب المستجوب في كل دراسة أن يعطي المعلومات التي تطلب منه بدقة كاملة ، وعقابه إذا رفض الاجابة أو إذا تبين انه اعطى اجابات خاطئة متعمدا . فاذا كان النص الأول يحث المستجوب على الاجابة فان هذا النص يلزمه بها وبذلك لا يكون هناك من مفر أمامه الا ان يعطي المعلومات التي تطلب منه . ولا يجب أن نفهم من ذلك ان وجود هذه النصوص في التشريع الاحصائي سوف يؤدي الى اجابة جميع المستجوبين على جميع الأسئلة

التي توجه اليهم ، ذلك لأن مديرية الاحصاء لا يمكن أن تلجأ الى تطبيق هذه النصوص في كل حالة رفض إذا انها لو فعلت ذلك سوف تظهر بمظهر الذي يقاضي أفراد الجمهور وبذلك تفقد تعاونهم المخلص الأمر الذي يدفعهم الى تضليلها بكل الوسائل التي يستطيعونها .

٣ - الزام مديرية الاحصاء بالمحافظة على سرية المعلومات التي تتجمع لديها وعقاب كل موظف يثبت افشائه لسرية هذه المعلومات .

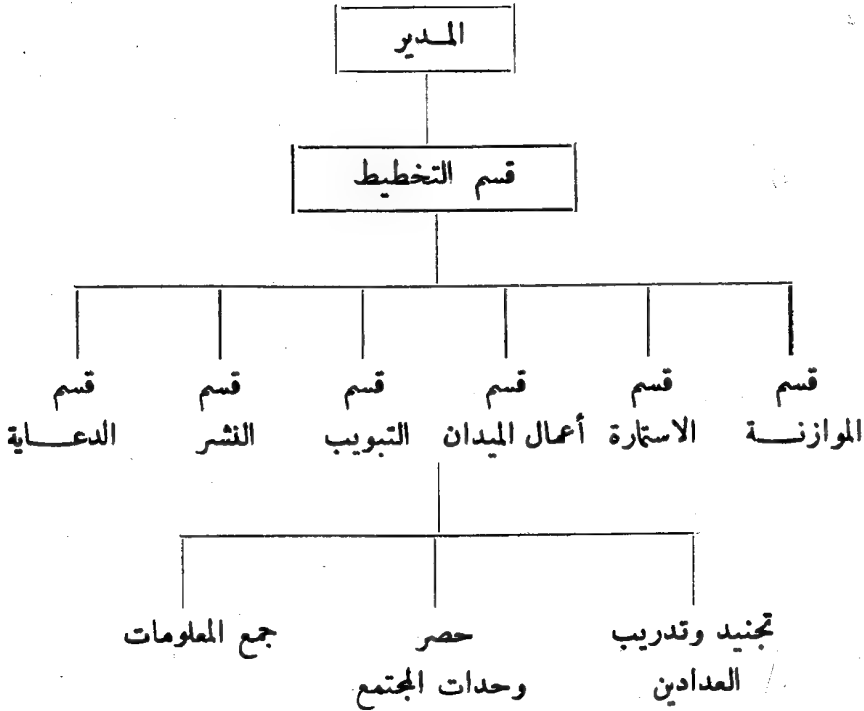
٤ - عدم استخدام المعلومات الاحصائية بشكل يعود بالضرر على صالح مستجوب معين أو مستجوبين معينين ، وبمعنى آخر لا تكون مديرية الاحصاء ملزمة بتقديم معلومات معينة تصل الى علمها عن مستجوب معين أو مستجوبين معينين الى أية جهة حكومية اخرى .

ويعني ذلك أن لمديرية الاحصاء الحق في رفض تزويد أية إدارة حكومية اخرى بمعلومات معينة عن مستجوب معين أو مستجوبين معينين، حيث أن واجب هذه المديرية هو تقديم بيانات احصائية مجمعة لا تظهر فيها أي معالم شخصية بتاتا . لذلك اذا وجدت مؤسسة واحدة تعمل في نشاط صناعي معين فان القانون في بعض الدول ينص على عدم نشر أية معلومات عنها الا بموافقتها الكتابية على ذلك، فاذا رفضت تلتزم مديرية الاحصاء بعدم النشر . ويسري هذا النص كذلك في حالة وجود مؤسستين فقط تعملان في نشاط صناعي معين حيث أن نشر بياناتهما وان كان في شكل ارقام مجمعة فيه افشاء لسرية بيانات كل منهما .

والهدف الاسامي من هذه النصوص هو طمأنة المستجوب وحفزه على اعطاء المعلومات التي تطلب منه صحيحة دون أي تضليل ، حيث لا معنى بتاتا لاتفاق الاموال على مديرية الاحصاء اذا كانت سوف تقوم بجمع معلومات خاطئة بسبب خوف المستجوبين وعدم اطمئنانهم .

والأساس الثاني الذي يجب أن يتوفر حتى يمكن القيام ببحث ميدان هو وجود جهاز من الموظفين المدربين على العمليات المختلفة التي يتطلبها مثل هذا البحث والمتفرغين لعملهم فقط . وفي بعض الدراسات التي تكون على نطاق واسع جواً (التعدادات) يحتاج العمل الى آلاف من الموظفين ولذا لا يمكن توظيفهم جميعاً على أساس دائم حيث ان ذلك يكلف الدولة أموالاً طائلة ولهذا يكون استخدام غالبيتهم على أساس مؤقت لمدة محدودة . والمشكلة التي يواجهها الجهاز القائم بالعمل في هذا الصدد هي في كيفية العثور على هذا العدد الكبير من الموظفين على أساس المستوى العلمي الذي يتحدد كشرط لتوظيفهم ، ولذلك يتقرر في الغالب اجراء عمليات الميدان في موسم الاجازات المدرسية حيث يكون هناك عدد كبير من الطلبة والمدرسين المتفرغين والذين بذلك يمكن الاعتماد عليهم . فاذا تكون الجهاز اللازم للقيام بالعمل لا بد من تقسيم الموظفين الذين سوف يوكل اليهم الاشراف على العمليات المختلفة الى مجموعات تتخصص كل مجموعة منها في عملية معينة بشرط أن لا تستقل برأيها استقلالاً تاماً حيث ان العمليات المختلفة متداخلة ومتفاعلة مع بعضها بحيث لا بد من تحقيق التعاون بين الاقسام المختلفة التي يتكون منها الجهاز . وبالطبع يجب أن تعمل هذه الاقسام وفقاً لما يرسمه قسم التخطيط الذي يكون مسؤولاً مع مدير الجهاز عن اتخاذ القرارات الأساسية التي يحتاجها العمل ، وعن تنظيمه والاشراف عليه بحيث يضمن انتهاء كل عملية في الموعد المحدد لها وبذلك يضمن انتهاء البحث في أقصر وقت ممكن دون الاخلال بالدقة التي تكون الهدف الأول من تنظيم العمل كما سبق أن ذكرنا . كذلك نلاحظ أن بعض الأقسام يكون عملها مركزياً فقط بينما يكون عمل البعض الآخر مركزياً ومحلياً في نفس الوقت ، ذلك انه في التعدادات الواسعة النطاق لا يمكن تجميع العدادين في مكان واحد لتدريبهم نظراً لضخامة عددهم فيكون بذلك من الأفضل تجميعهم على أساس محلي . كما أن الاشراف عليهم في الميدان يكون أكثر فعالية عندما تكون صلة العدادين بالمشرفين عليهم صلة مباشرة في المنطقة التي يعملون

فيها . والرسم التخطيطي التالي يوضح الاقسام المختلفة التي يتكون منها الجهاز الذي يجب تكوينه للقيام بتعداد على نطاق واسع .



والأساس الثالث الذي يجب توفره حتى يمكن ان يقوم الجهاز بعمله على الوجه الأكمل هو تخصيص الاموال التي يتطلبها العمل واعطاء مدير الجهاز سلطة انفاقها دون التقيد بالروتين الحكومي ولكن في حدود الانظمة الخاصة بالصرف . ان عدم تخصيص الاموال الكافية والتقيد بالروتين الخاص بالانفاق يعطل الجهاز في عمله ويجعله تحت رحمة المسؤولين ، الأمر الذي يمكن أن يؤدي إلى فشل العملية بكاملها حيث ان العمل يجري تبعاً لجدول زمني وعدم التقيد به بسبب عجز الاموال أو بسبب الروتين يؤدي الى ارتباك . فضلاً عن ذلك تكون مديرية الاحصاء قد ارتبطت بموعد معين لاجراء عملية العد أمام

جمهور المستجوبين ، وارثباك العمل يؤدي الى عدم امكانية التقيد بهذا الموعد ، الأمر الذي يؤدي إلى فقدان الثقة في اعمالها وتدهور هيبتها . ويرتبط بتخصيص الأموال اللازمة تزويد الجهاز القائم بالعمل بالآلات الاحصائية التي بدونها لا يمكن انجاز عمليات التبويب في الوقت المناسب خاصة في التعدادات الواسعة النطاق .

فاذا توفرت هذه الأسس الثلاث يستطيع الجهاز القيام بعمله متتبعاً الخطوات الآتية ، على انه يجب أن نلاحظ أن تسلسل هذه الخطوات لا يعنى تسلسلها فعلاً في الواقع العملي حيث سبق أن ذكرنا أن العمل يقوم على أساس توزيعه بين مجموعات من الموظفين يمكن أن يقوم كل مجموعة منها بالعمل الذي يوكل اليها في نفس الوقت الذي تقوم فيه المجموعات الأخرى بعملها . وفضلاً عن ذلك فان طبيعة العمل تستدعي أحياناً اتخاذ قرارات بشأن عدة عمليات مجتمعة في وقت واحد .

أولاً : تصميم اطار البحث وخطته العامة :

وبستدعى ذلك القيام بعدد من العمليات نردها فيما يلي : -

١ - تحديد الهدف او الاهداف من البحث : ويعني ذلك تحديد النقاط المختلفة التي سوف يدور حولها البحث . وتحديد أهداف البحث ليس عملاً احصائياً بحتاً فلا بد أن يكون لدى الباحث ادراك واضح وفهم سليم لمدى الترابط بين هذه الأهداف ، اي لا بد أن يكون ملماً علماً صحيحاً بموضوع البحث .

ان تحديد أهداف البحث يفيد أولاً في تحديد المعلومات المختلفة التي سوف يسأل عنها ، فلو كان أحد أهدافنا من بحث عن الصناعة ، مثلاً ، تحديد القيمة الصافية المضافة في هذا النشاط ، فان استمارة البحث لا بد وان تشمل

على الأسئلة المختلفة التي يمكن بواسطتها التوصل إلى حساب هذه القيمة ، أما إذا كان تحديد هذه القيمة ليس من أهدافنا فلا داعي لوجود مثل هذه الأسئلة . ويجب أن نتذكر دائماً أن استمارة البحث كلما طالت واشتملت على أسئلة كثيرة كلما أدى ذلك إلى صعوبات كثيرة عند العمل في الميدان وفي المكتب (التبويب) وكلما أدى ذلك أيضاً إلى مضايقة المستجوبين ، الأمر الذي قد يدفعهم إلى رفض الإجابة أو التخلص منها بإعطاء اجابات سريعة غير حقيقية . لذلك يجب ان يعمل الباحث كلما كان ذلك ممكناً على الا تزيد اسئلة الاستمارة كثيراً ، وتحديد أهدافه من بادىء الأمر يساعده في عدم توجيه أسئلة لا يكون البحث في حاجة إليها .

كذلك يساعد تحديد أهداف البحث في تحديد المجتمع الإحصائي موضوع البحث أي في تحديد مجموع الوحدات التي سوف تعتبر وحدات استبيان عند النزول إلى الميدان لجمع المعلومات . فإذا كنا نريد القيام ببحث عن ميزانية الأسرة ، مثلاً ، وكان هدفنا من البحث استنتاج الاوزان التي ترجح بها أسعار السلع التي تدخل في تركيب الرقم القياسي لنفقة معيشة الطبقة العاملة (سنناقش ذلك بتفصيل في موضوع الارقام القياسية) فان مجتمع البحث يتحدد تبعاً لذلك بأسر الطبقة العاملة التي لا يزيد دخلها الشهري عن مبلغ معين ، أما إذا كان هدفنا من البحث التوصل إلى تقدير استهلاك المجتمع عامة من السلع المختلفة فان مجتمع البحث في هذه الحالة لا بد أن يشتمل الأسر عامة بصرف النظر عن مستوى دخلها . كذلك في الحالة الأولى فان أسئلتنا سوف توجه إلى قيم الانفاق على السلع المختلفة ، أما في الحالة الثانية سوف تكون الاسئلة عن الكميات المستهلكة من السلع المختلفة في مدة معينة .

وفي الأبحاث التجريبية تتخذ هذه الخطوة شكلاً آخر حيث يتعين على الباحث تحديد وتعريف المشكلة التي سوف يدور حولها البحث . وفي هذه

الحالة أيضاً يؤدي تحديد وتعريف المشكلة الى تحديد المعلومات التي سوف يسأل عنها والمجتمع الاحصائي الذي سوف تجرى عليه التجربة .

وقبل تحديد المعلومات التي يريد الباحث ان يسأل عنها يجب دراسة كل المراجع الممكنة للوقوف على ما جمع من هذه المعلومات فعلاً في دراسات سابقة ، حيث ان ذلك يساعد في اختصار استمارة البحث وهذا يوفر كثيراً من الجهود والتكاليف .

٢ - تحديد نطاق وشمول البحث ، أي تحديد وتعريف المجتمع موضوع البحث ويساعد في ذلك كما قدمنا تحديد الأهداف . على انه لا يكفي تحديد المجتمع بشكل عام بل يجب تعريفه تعريفاً دقيقاً ويتم ذلك بتحديد وتعريف الوحدة التي يتكون منها هذا المجتمع . ففي بحث عن الصناعة يكون المجتمع الاحصائي هو مجموع المؤسسات الصناعية في الدولة ، ولكن ما هو تعريف المؤسسة الصناعية الذي يحدد لنا هذا المجتمع تحديداً دقيقاً ، هل هي المؤسسة التي تقوم بالانتاج الصناعي أو الاصلاح أو الصيانة في مكان محدد وقت جمع المعلومات ، أم هي المؤسسة التي تقوم بالانتاج الصناعي فقط وبذلك لا يدخل في المجتمع الذي نبهته المؤسسات التي يقتصر نشاطها على الاصلاح والصيانة فقط . واذا كنا نقوم ببحث عن الزراعة يكون المجتمع الاحصائي هو مجموع الحيازات الزراعية في الدولة ، ولكن هل يدخل في هذا المجتمع مجموع الحيازات بشكل عام مهما كانت صغيرة أو كبيرة أو سنقتصر فقط على الحيازات التي تزيد مساحتها عن مساحة معينة . وكذلك في دراسة عن السكان هل يدخل في تعريفنا المجتمع البحث السكان الرحل أيضاً ، أم سوف نستبعدهم .

ومن الواضح أن وضع تعريف معين لوحدة المجتمع الاحصائي الذي نقرر دراسته لا يعني اننا سوف نجد ان هذا التعريف ينطبق في الواقع العملي انطباقاً كاملاً ، حيث انه في الغالب نصادف مشاكل معينة عند ربط التعريف

النظري الذي نستخدمه بالواقع العملي . ومهمة الباحث هي توقع هذه المشاكل وإيجاد حل لها حتى يكون العداد في عمله في الميدان على بينة من أمره . كما أن الباحث يجب أن ينبه العدادين الى أنهم قد يصادفون مشاكل أخرى غير تلك التي نبههم اليهم ، فإذا حدث ذلك لا يجب أن يتصرفوا الا بعد أخذ رأي القائمين على الاشراف عليهم ، وبهذه الطريقة يمكن حل المشاكل المختلفة التي تصادف العدادين بالنسبة لتعريف وحدة الاستبيان حلاً موحداً . فإذا كان بحثنا عن الصناعة واستخدمنا التعريف الذي سبق الإشارة اليه للمؤسسة الصناعية فسوف نجد في الواقع العملي مؤسسات تعمل في أنشطة اقتصادية مختلفة ، مثلاً مؤسسة تربي الماشية وتأخذ منها الحليب فتصنعه ثم تباعه الى المستهلكين ، هذه المؤسسة تعمل في النشاط الزراعي والصناعي والتجاري ، هل تدخل في المجتمع الإحصائي موضوع البحث أو تستبعد ؟ وإذا ادخلناها كيف يمكن جمع المعلومات الخاصة بنشاطها الصناعي فقط ؟ هذه الأسئلة وغيرها تعبر عن المشاكل العملية التي تظهر عند تطبيق تعريف وحدة الاستبيان وواجب الباحث أن يضع لها الحلول حتى يكون عمل العدادين جميعاً على أساس موحد .

٣ - هل يكون البحث آنياً أو على مراحل ؟ والمقصود بالبحث الآني

توحيد الفترة الزمنية التي تجمع فيها المعلومات من جميع وحدات المجتمع الإحصائي الذي تقرر دراسته . ومن الواضح ان إجراء العمل بهذه الطريقة يستدعي استخدام عدداً كبيراً جداً من العدادين ، الأمر الذي يوقع الجهاز القائم بالعمل في مشاكل كثيرة ، منها الاضطرار إلى استخدام عدادين من مستوى علمي متدني ، وصعوبة تدريبهم التدريب الكافي ، وصعوبة الاشراف عليهم في الميدان ، وصعوبة تنظيم العمل الميداني نفسه في الفترة المحددة لجمع المعلومات ، ثم استقبال المكتب لعدد كبير جداً من الاستمارات بعد الانتهاء منها في الميدان في وقت واحد ، الأمر الذي يؤدي الى ارباك العمليات

المكتبية المختلفة (المراجعة والرميز الخ) . ولا شك ان الجهاز يحاول معالجة هذه الصعوبات بما يراه مناسباً ، الا أنه منها كان ماهراً في ذلك لا بد وان يترتب على العمل بهذه الطريقة أخطاء كثيرة تنقص من دقة النتائج الأخيرة . وتلعب هذه الطريقة عندما يكون المجتمع الاحصائي موضوع البحث مجتمعاً متحركاً أي عندما تكون معلوماته قابلة للتغيرات السريعة مثل المجتمع السكاني ، فالسكان في حركة دائبة تشمل مواليد جدد ووفيات وتغيرات مستمرة في الحالة الزوجية والتعليمية ... الخ ولهذا يضطر الجهاز القائم بتعداد السكان إلى أن يجري العمل بهذه الطريقة بالرغم من معرفته للصعوبات المختلفة التي سوف تواجهه معتمداً على الوسائل المختلفة التي يحاول اتباعها للتخفيف من حدة هذه الصعوبات ولانقاص الأخطاء التي يمكن ان تترتب عليها .

أما إذا كان المجتمع موضوع البحث لا يتعرض لهذه الحركة الدائبة يكون من الأفضل اجراء العمل على مراحل . وذلك بتقسيم المجتمع إلى أقسام والبدء بجمع المعلومات من قسم معين أولاً ثم ينتقل فريق العدادين المكلفين بالعمل الى الأقسام الاخرى تباعاً . ان العمل بهذه الطريقة يقضي على كثير من الصعوبات التي سبق الإشارة اليها ، حيث يكون العمل في حاجة الى عدد قليل من العدادين يمكن انتقاؤهم على أساس المستوى العلمي اللازم للبحث وعلى أساس شخصية العداد ولباقة وبذلك يمكن التثبت من عنصر هام من عناصر نجاح العمل ، كذلك يكون تنظيم العمل الميداني والاشراف عليه أسهل بكثير من البحث الآني ، كما ان الاستمارات التي تنتهي في الميدان ترد إلى المكتب تباعاً وعلى مراحل الأمر الذي يخفف من ضغطها ، وبالإضافة الى ذلك يزداد العدادون خبرة بعملهم كلما انتقلوا من قسم من المجتمع إلى قسم آخر .

ويرتبط بهذا القرار كذلك تحديد موعد النزول إلى الميدان لجمع المعلومات وهل تجمع المعلومات مرة واحدة أو على فترات دورية . هناك عوامل كثيرة

يجب أن تؤخذ في الاعتبار عند تحديد موعد النزول إلى الميدان لجمع المعلومات ، منها مدى الاحتياج إلى العدادين المتفرغين فكلما زاد الاحتياج كلما أصبح من الواجب أن يكون موعد جمع المعلومات خلال العطلة المدرسية الصيفية . كما ان ظروف الطقس يمكن أن يكون لها تأثيرها خاصة في الدراسات الزراعية إذ لا يجب النزول إلى الميدان في موسم الامطار وبشكل خاص عندما تكون طرق المواصلات في المناطق الزراعية في حالة رديئة ، كذلك لا يجب النزول إلى الميدان في الدراسات السكانية في المواسم التي يكون السكان فيها في حركة شاذة بسبب الاصطياف مثلا . وفي البحوث الخاصة بميزانيه الاسرة يجب جمع المعلومات على فترات دورية حتى يمكن جمع معلومات تمثل التغيرات الموسمية في انفاق الاسر . كذلك في البحوث التي يكون الهدف منها معرفة التغيرات الجارية في الظواهر موضوع البحث مثل التوظيف والبطالة والاجور والانتاج الصناعي يكون من الواجب جمع المعلومات في فترات دورية .

٤ - تحديد طريقة جمع المعلومات - وتجمع المعلومات عادة باحدى

طريقتين أساسيتين ، هما طريقة الادلاء الشفهي وطريقة التسجيل الذاتي . وفي الطريقة الاولى يقوم العدادون بالمقابلة الشخصية للمستجوبين ويوجهون اليهم الأسئلة ثم يسجلون إجاباتهم في استمارة البحث . والاستمارة في هذه الطريقة تسمى كشف البحث . أما الطريقة الثانية فتقوم على أساس الاتصال غير المباشر بالمستجوبين وذلك بواسطة البريد أو التليفون . والاستمارة التي تستخدم في هذه الطريقة تسمى صحيفة الاستبيان .

على ان العمل يجري في الغالب بالجمع بين الطريقتين حيث يقوم العدادون بتوزيع الاستمارات على المستجوبين ويطالبونهم بملأها إذا تبينوا منهم القدرة على القراءة والكتابة إلى الدرجة التي تمكنهم من القيام بذلك ، ثم يزورونهم مرة ثانية لجمع الاستمارات ومساعدة المستجوبين في اجابة الأسئلة التي لم

يتمكنوا من فهمها أو التي لم يتمكنوا من فهم كيفية كتابة الاجابة عليها . كذلك يمكن أن ترسل الاستمارات إلى المستجوبين بالبريد وبعد ورود الاستمارات الى المكتب يمكن تحديد المستجوبين الذين لم يهتموا باعادة استماراتهم فيقوم العدادون بزيارة هؤلاء المستجوبين للحصول على إجاباتهم ، وبذلك نستطيع أن نقول أن الجهاز القائم بالعمل لا يجب أن يتقيد بطريقة معينة تقيداً جامداً ، إذ يجب أن يتبع ما تطلبه عليه طبيعة البحث وتطوره .

وبالنسبة لبعض البحوث تكون الطريقة التي يجب اتباعها هي الملاحظة حيث يقوم العدادون بملاحظة الظاهرة موضوع البحث ثم كتابة المعلومات دون توجيه الأسئلة كما هو الحال في دراسة عن كيفية استجابة الاطفال في سن معينة مثلاً لأنواع معينة من المؤثرات ، وكذلك في دراسة عن حركة السير حيث يقوم العدادون بملاحظة السيارات ثم تدوين ملاحظاتهم في الاستمارة المعدة لذلك . وبالنسبة لبعض البحوث الأخرى تكون الطريقة التي يجب اتباعها هي التسجيل وذلك في الحالات التي تكون فيها الهيئة المشرفة على البحث لها من السلطة التي تعطيها الحق في الزام المبحوثين بتقديم البيانات المطلوبة لتسجيلها في سجلات معينة (سجلات الجمارك - سجلات المواليد - سجلات الموظفين .. الخ) .

ولكل من كشف البحث وصحيفة الاستبيان مزاياه وعيوبه ، وتعتبر مزايا كل منها عيوباً للآخر والعكس . ومن أهم مزايا كشف البحث ما يلي :

١ - يساعد كشف البحث على جمع بيانات بالملاحظة المباشرة لبعض الظواهر دون الحاجة الى توجيه اسئلة عنها خاصة اذا كانت هذه الاسئلة منفرة أو محرجة ولا يحتمل الحصول على اجابات دقيقة عليها فمثلاً يستطيع العداد عند زيارته لاسرة ما ان يشاهد حالة أثاث المنزل ووسائل الاضاءة وملابس الأفراد وغير ذلك من البيانات التي لا يمكن السؤال عنها لما قد يكون في ذلك

من أثره في نفس المستوجب قد يجعله يرفض الاجابة على باقي الاسئلة أو يعطي اجابات غير دقيقة .

٢ - تصلح هذه الوسيلة في الدراسات التي تتطلب اسئلة كثيرة العدد، اذ ان العدد يستطيع بلباقته أن يكسب معونة المستجوبين باقناعهم بما لهذه الدراسة من نتائج تعود على المجتمع بالنفع والفائدة ، كما يستطيع أن يسهل مهمتهم بشرح الاسئلة بما يتناسب وادراك كل منهم وصرف النظر عن الاسئلة التي يبدي المستجوب عدم الرغبة في الاجابة عليها .

٣ - ليس من المستحب ولا من المتيسر أن يكلف المستجوبين بقراءة تعليقات مطولة عن معاني الاسئلة وكيفية الاجابة عليها اذ قد لا يكون لديهم الوقت الكافي او الاستعداد لقراءة هذه التعليقات . لهذا يعتبر كشف البحث من انجح الوسائل لجمع المعلومات في حالة الدراسات التي تتطلب توجيه اسئلة يكون من الصعب الاجابة عليها دون شرح خاص حيث ان العدد يمكن أن يقوم بتوضيح معاني الكلمات وشرح المقصود من الاسئلة كلما استدعى الامر ذلك .

٤ - من الملاحظ أن الشخص لا يجيب على السؤال الذي يوجه اليه إلا تبعاً للمعنى الذي يفهمه من هذا السؤال . فاذا فهم معنى يخالف المقصود فعلاً من السؤال كانت اجابته عديمة القيمة بل قد تؤدي الى نتائج مضللة وغير دقيقة . ومن النادر ان يحدث ذلك في حالة كشف البحث ، اذ انه اذا اتضح للعدد ان المستجوب أعطى اجابة غير متفقة مع المعنى المقصود من السؤال فانه يقوم بشرحه وتوضيحه حتى يحصل على الاجابة الدقيقة له .

٥ - اذا شك العدد في صحة الاجابات التي يحصل عليها أو إذا اتضح له تناقض الاجابات تناقضاً ظاهراً فانه يستطيع مناقشة المستجوب حتى يحصل على بيانات متأسكة لا تناقض فيها .

٦ - باستخدام كشف البحث كأداة لجمع المعلومات يكون من المؤكد إلى درجة بعيدة الحصول على اجابات لكل الاسئلة من جميع المستجوبين (أو على

الأقل من نسبة كبيرة منهم) خاصة إذا تحققت جميع العوامل الأخرى التي تساعد على نجاح الدراسة مثل اختيار العدادين بحيث يصلحون لهذا العمل من جميع النواحي وتدريبهم التدريب الكافي على كيفية اجراء العمل بدقة والقيام بالدعاية اللازمة لجميع الطرق الممكنة في حالة الدراسات الواسعة النطاق ، وإذا أحسن العدادون عرض أغراض الدراسة على المستجوبين واقناعهم بأهمية تعاونهم ، وإذا اختير الوقت المناسب للاتصال بالمستجوبين .

إلا أن كشف البحث له عيوب أهمها : -

١ - كثرة التكاليف والمجهودات التي تبذلها الهيئة المشرفة على الدراسة في اختيار العدادين وتدريبهم ودفع مرتباتهم واجور انتقاهم .

٣ - يكون للعدادين سلطة كبيرة على الاجابات التي تؤخذ من المستجوبين وحتى يمكن أن تأتي هذه الاجابات صحيحة ودقيقة يجب أن يتصف العدادون بالحياد التام ، أما إذا كان بعض العدادين متحيزين لافكار معينة (خاصة في الدراسات الاجتماعية) فان اتصاهم الشخصي بالمستجوبين يتيح لهم فرصة التأثير عليهم وتوجيه الأسئلة اليهم بحيث تأتي الاجابات مؤيدة لافكارهم . وهذا بالطبع يتنافى مع الروح العلمية التي يجب أن تتصف بها الدراسة .

٣ - في بعض الدراسات تشتمل استمارة البحث على أسئلة قد تضر بالمستجوب الذي يدلي بها ضرراً مادياً أو أدبياً . وذلك مثل الاسئلة عن تعاطي المخدرات أو العلاقات الجنسية أو اعتناق المبادئ السياسية التي تحرمها الدولة وما شابه ذلك ، وفي هذه الدراسات لا يمكن أن يصلح كشف البحث في الحصول على اجابات لاسئلة من هذا النوع نظراً للمقابلة الشخصية بين العداد والمستجوب .

أما صحيفة الاستبيان فلا تختلف كثيراً عن كشف البحث إلا في طريقة ملئها فهي اما أن تسلم للمستجوب أو ترسل اليه بالبريد أو تنشر على صفحات

الجرائد والمجلات (في بعض الدراسات) ويقوم المستجوب بتدوين اجاباته على اسئلتها ثم يعيدها إلى الهبة المشرفة على الدراسة . ولصحيفة الاستبيان كما لكشف البحث مزايا وعيوب أهمها من ناحية المزايا ما يلي :

١ - قلة التكاليف والمجهودات اللازمة لجمع البيانات خاصة إذا أرسلت بالبريد أو نشرت على صفحات الجرائد حيث يمكن توفير الغناء في تجنيد العدادين وتدريبهم ودفع اجورهم وغير ذلك من نفقات .

٢ - تصلح في الدراسات التي تتطلب الحصول على معلومات لا يمكن الجهر بها لما قد ينتج من ذلك من ضرر أو لأن التقاليد لا تسمح بالتصريح بمثل هذه البيانات . حيث باستخدام صحيفة الاستبيان يمكن عدم كتابة اسم المستجوب وبذلك لا يمكن التعرف على صاحبها الأمر الذي يشجعه على اعطاء الاجابات المطلوبة بصراحة تامة .

٣ - كما انه باستخدام صحيفة الاستبيان لا تكون الدراسة تحت رحمة العدادين من ناحية امالهم أو تحيزهم .

أما عيوب صحيفة الاستبيان فيمكن تلخيصها في الآتي :

١ - لا يمكن استخدامها إلا إذا كان المستجوبون مثقفين أو على الأقل يجيدون القراءة والكتابة .

٢ - يكون تصميم صحيفة الاستبيان من الصعوبة بمكان كبير حيث يجب أن تكون الاسئلة في غاية البساطة والوضوح ولا تحتل التأويل حيث لن يكون هناك عدادون يساعدون المستجوبين في فهم الاسئلة الفهم الصحيح .

٣ - لا يمكن التوسع في الاسئلة الموجهة الى المستجوبين لأن ذلك يمكن أن يؤدي الى امال الاجابة عليها .

٤ - إذا كانت الدراسة تستدعي عدم كتابة اسم المستجوب للتشجيعه

على اعطاء الاجابات الصحيحة ، فانه لا يمكن الرجوع اليه لاستكمال ما قد يكون بالصحيفة من نقص أو لتصحيح الاخطاء أو الاستفسار عن الاجابات المتناقضة ان وجدت .

٥ - قلة عدد صحائف الاستبيان التي ترد إلى الهيئة المشرفة بعد ملئها بها وضعت هذه الهيئة من مغريات وتسهيلات للاجابة خاصة إذا كانت الصحائف ستعاد بالبريد (هذا بالمقارنة مع كشف البحث) . ولذلك أهمية كبيرة خاصة إذا كانت الدراسة تجري بالمعينة حيث يصبح عدد الاجابات التي تحصل عليها الهيئة المشرفة على الدراسة غير متفقاً مع حجم العينة الذي حدد أولاً ، كما ان الصحائف التي تعاد إلى الهيئة لا يمكن اعتبارها ممثلة للمجتمع موضوع الدراسة ، الأمر الذي يدخل عنصر التحيز في النتائج واحياناً يتضخم هذا العنصر فتفشل الدراسة . ويكون هذا التحيز راجعاً الى اهتمام البعض بموضوع الدراسة لسبب معين ، مثلاً ، محاولة استخدام الدراسة كفرصة لاطهار رأيهم وعدم اهتمام البعض الآخر ، الأمر الذي يدفعهم إلى عدم الاجابة . ولعل ذلك هو السبب في فشل الدراسات الخاصة بالاستقصاء عن نتائج الانتخابات في بعض الدول ، حيث يكون المؤيدين لمرشح ما واثقين من نجاحه فلا يهتمون بالاجابة على الاستقصاء ، بينما يتخذ المعارضون الاستقصاء وسيلة للتعبير عن اتجاههم فيجيبون ، ولهذا تأتي النتيجة غير متفقة مع ما يحدث فعلاً .

٦ - أما بالنسبة للصحائف التي تعاد فعلاً الى الهيئة المشرفة على الدراسة فكثير منها سوف يشتمل على اجابات تهكمية أو مفترضة ، وبالطبع لن تكون لهذه الصحائف أي قيمة بل ستؤدي إلى ضياع الوقت في قراءتها .

على انه يمكن استخدام بعض الأساليب والوسائل إلى التقليل من أثر العيوب السابقة ومن أهم هذه الأساليب العمل على توضيح أغراض الدراسة توصيماً جذاباً وبكل الوسائل الممكنة (خاصة في الدراسة الواسعة النطاق)

بحيث تظهر الفوائد العملية والعلمية لنتائجها الدقيقة ، وكذلك يمكن أن يرفق بالصحيفة عند إرسالها للمستجوب خطاب رقيق يوضح أغراض الدراسة بأمانة ودقة ويبين للمستجوب أهمية تعاونه بإعادة الصحيفة بعد ملئها والفائدة التي تتحقق بإعطائه الاجابات الصحيحة والدقيقة . ولتسهيل مهمة المستجوب اذا طلب منه اعادة الصحيفة بالبريد يحسن أن يرسل اليه مع الصحيفة مغلف عليه طابع بريد وعنوان الهيئة المشرفة على الدراسة ، ويا حبذا لو تضمن المغلف على أكثر من نسخة من الصحيفة لمواجهة احتمال تلف احداها .

٥ - تحديد مجال جمع المعلومات ، أي هل سيشمل جمع المعلومات جميع وحدات المجتمع الاحصائي الذي تقرر بحثه (العدد الشامل) ، أو عينة من هذا المجتمع ثم تعميم نتائجها على المجتمع الأصلي ؟ ولا شك أن هناك اسباب تؤدي الى هذا السؤال ، الأمر الذي يقودنا الى مناقشة مزايا الدراسة بالمعينة .

يؤدي استخدام المعينة الى توفير جزءاً كبيراً من التكاليف والجهود التي تضطر اليها في حالة العد الشامل حيث اننا سنستخدم جزءاً من المجتمع وليس كله . على انه يجب أن نلاحظ ان توفير التكاليف لا يكون بنسبة حجة العينة الى حجم المجتمع الذي سحبت منه ، فاذا كان حجم العينة يكون ١٠٪ من حجم المجتمع فلا يترتب على ذلك انخفاض التكاليف الى العشر حيث ان الدراسة بالمعينة تستدعي اجراء خطوات عمل وخبرة رياضية لا يحتاجها العد الشامل ، الأمر الذي يوجد عناصر جديدة من التكاليف . ولهذه الميزة أهمية كبيرة حيث أن كثيراً من البحوث لو لم يكن من الممكن اجرائها بالمعينة لما قامت بها الهيئات المعنية نظراً لارتفاع التكاليف التي لا قبل لها بتحملها اذا كان لا بد من اجرائها بالعد الشامل ، ولعل ذلك يفسر كثرة البحوث الاحصائية في الوقت الحاضر وخاصة في الدول التي لا تستطيع أن تتحمل ميزانياتها نفقات العد الشامل الباهظة . كما ان لتوفير الجهود أهمية

كبيرة حيث ان المستوى التعليمي في بعض الدول لا يسمح لها بتزويد الاجهزة الاحصائية بالعدد اللازم من الموظفين الذين يحتاجهم إذا كان يتحتم عليها اجراء جميع الدراسات الاحصائية بالعد الشامل . وبشكل عام نستطيع أن نقول أن امكانية اجراء البحوث الاحصائية بالمعينة فتحت مجالات واسعة للبحث لا يمكن من الممكن تصورها من قبل .

كذلك يؤدي استخدام المعينة الى توفير الوقت الذي يضيع في اجراء البحوث بالعد الشامل ، ولتوفير الوقت أهمية كبيرة ، ففي كثير من الدراسات يؤدي اتساع نطاق العمل بالعد الشامل الى ظهور النتائج بعد سنوات طويلة (تعداد السكان مثلا) بحيث تصبح هذه النتائج صورة لما كان عليه المجتمع قبل بضع سنوات ، أي تصبح مجرد صورة تاريخية لا يكون لها أهمية كبيرة في اتخاذ اجراءات عملية على أساسها (التخطيط الاقتصادي والاجتماعي مثلا) . وفي بعض الحالات تكون الدولة في حاجة الى بيانات احصائية بعد وقت قصير حتى تستطيع أن تقرر سياستها نحو الظاهرة المتعلقة بهذه البيانات ، مثلا عندما ترغب الحكومة في تقدير محصول القمح في عام معين حتى إذا تبينت عجزه بدأت تعمل على استيراد ما تحتاجه من الخارج والعكس إذا تبينت زيادته عما يكفي المجتمع تبدأ في محاولة تصريف الفائض منه في الأسواق الخارجية . وفي مثل هذه الحالة يتحتم اجراء البحث بالمعينة نظراً لأن اجرائه بالعد الشامل لا يتفق مطلقاً مع الغرض منه .

كذلك يمكن أن يكون البحث أكثر تفصيلاً في الدراسات بالمعينة ، حيث انه في حالة اجرائه بالعد الشامل يأخذ الباحث اتساع نطاق البحث في اعتباره فيحاول دائماً انقاص عدد الاسئلة في الاستارة لأن زيادتها يزيد المشكلة تعقيداً وتكاليفاً وجهداً . ان اجراء البحث بشكل تفصيلي يحتاج في الغالب إلى عدادين اكفاء متخصصين ومدربين تدريباً عالياً على العمل الذي يوكل اليهم . ومن الواضح انه لا يمكن أبداً الحصول على هذا النوع من العدادين في الدراسات

بالعد الشامل . واذا أخذنا في اعتبارنا جميع خطوات العمل من حيث الوقت والجهد والتعقيد أدركنا انه يستحيل اجراء البحوث بشكل تفصيلي بواسطة العد الشامل . لذلك تساعدنا الدراسة بالمعينة على اجراء بحوث أكثر تفصيلاً وأكثر ارتباطاً بالواقع تبعاً لذلك ، الأمر الذي يزيد من اهميتها العلمية . وقد أدركت الاجهزة الاحصائية في كثير من الدول هذه الميزة فعملت على اجراء دراسات بالمعينة بجانب تعداد السكان الواسع النطاق وأثناء القيام به للتعرف على الصفات التفصيلية للسكان لمساعدة الحكومة في تخطيط الشؤون المختلفة الخاصة بهم .

كذلك يمكن إنقاص الأخطاء العامة التي سبق الاشارة اليها إلى أدنى حد ممكن بسبب ضيق نطاق العمل . حقاً ان الدراسات بالمعينة تؤدي الى ظهور نوع آخر من الخطأ لا يظهر في العد الشامل وهو الخطأ الكامن في عملية المعينة نفسها ، حيث نلاحظ انه مهما دققنا في اختيار وحدات العينة ومهما اتبعنا في ذلك من طرق صحيحة فان العينة الناتجة لا يمكن ان تكون صورة طبق الأصل من المجتمع الذي سحبت منه ، وبمعنى آخر لا يمكن ان نقرر ان المقاييس التي نصل اليها عن طريق المعينة تكون هي ذاتها المقاييس (المتوسط مثلاً) التي كان يمكن أن نصل اليها لو أجرينا الدراسة بالعد الشامل ، فلا بد أن يوجد فرق وهو الذي نسميه خطأ المعينة ، وينشأ هذا الخطأ بسبب اختلاف وزن الأنواع المختلفة من الوحدات في كل من المجتمع والعينة ، وغالباً تكون الوحدات الشاذة ذات وزن قليل في المجتمع نظراً لضخامة عدد وحداته وذات وزن اكبر في العينة نظراً لقلّة عدد وحداتها . على ان النظرية الرياضية الخاصة بالاحتمالات تساعدنا في حساب هذا الخطأ وهو أمر من الأهمية بمكان كبير حيث انه اذا لم يكن من الممكن حساب هذا الخطأ لكانت الدراسة بالمعينة عديمة الأهمية بتاتاً ، فنحن نعلم في هذه الحالة بوجود فرق بين المقاييس الناتجة من العينة وتلك التي تنتج من العد الشامل ولكننا لا نعلم كم هو هذا الفرق ، الأمر الذي يجعل دراستنا عمياء لا قيمة لها .

وبذلك إذا افترضنا مثلاً ان الأخطاء العامة في العد الشامل تصل الى ١٠ ٪ ، وأمكن إنقاصها بالمعينة الى ٢ ٪ ، وإذا افترضنا ان خطأ المعينة هو ١ ٪ يكون مجموع الخطأ في نتائج المعينة هو ٣ ٪ فقط بالمقارنة مع أخطاء العد الشامل التي تصل الى ١٠ ٪ ؛ لذلك تكون الميزة الأساسية للأبحاث بالمعينة هي زيادة دقة النتائج خاصة في الدراسات التي تدور حول موضوعات معقدة ، إذ انه في هذه الدراسات يكون هناك مجال واسع لانقاص الأخطاء العامة حيث نستطيع أن نستفيد من ارتفاع مستوى العدادين ومن اعطائهم الوقت الكافي لشرح كل سؤال شرحاً جيداً للمستجوبين .

ولقد لاحظنا مثلاً أن أخطاء الاجابة يمكن القضاء عليها اذا امكن جمع بيانات موضوعية بالقياس الفعلي للظواهر موضوع البحث (قياس مساحات الأراضي الزراعية مثلاً) ومن الواضح انه لا يمكن استخدام هذه الطريقة في حالة الدراسات بالعد الشامل حيث لا يكون هناك الوقت الكافي لذلك .

كذلك هناك مجتمعات يكون من المستحيل بأي حال من الأحوال اجراء بحثها احصائياً بالعد الشامل مثل الأممك والطيور والحيوانات المفترسة الخ ، وبذلك لا يكون هناك مفرأ من استخدام المعينة ، كذلك لا يجب أن ننسى ان التجارب العلمية ما هي الا دراسة بالمعينة .

كذلك يمكن أن يتبين لنا بعد اجراء البحث بالعد الشامل ان المستجوبين تحيزوا في اجابة سؤال معين وليس هناك وسيلة لتصحيح هذه الاجابات الا بأخذ عينة منهم ومعاودة توجيه السؤال بروية واهتمام أكبر ، ولعل في ذلك ما يساعدنا على الحصول على نتائج أكثر دقة .

لهذه الأسباب مجتمعة انتشر استخدام العينات في مختلف ميادين البحث العلمي واستعاض بها عن العد الشامل ، الأمر الذي دفع علماء الاحصاء إلى تطوير الأبحاث الخاصة بهذه الطريقة في البحث حتى أصبحت تكون فرعاً مستقلاً في علم الاحصاء . على ان المزايا الكثيرة التي تترتب على استخدام

المعينة لا تعنى ان اتباعها لا يواجهنا بأية صعوبات ، ذلك لأن هنالك مبادئ عامة اساسية يجب ان نأخذها دائماً بعين الاعتبار عند بحث هذا الموضوع ويمكن تلخيصها في الآتي :

١ - ان سحب العينة التي يمكن أن تكون صورة صادقة للمجتمع الذي سحبت منه يستلزم معرفة بعض خصائص هذا المجتمع مقدماً حتى يمكن تحديد نوع العينة التي تناسبه وحجم العينة التي يجب سحبها والطريقة التي يجب اتباعها في عملية السحب وبدون هذه المعرفة تصبح المعينة أمراً متعذراً مستحيلاً . ان المجتمعات التي يمكن أن نرغب في بحثها احصائياً ليست جميعاً من نفس النوع ، ففي بعض هذه المجتمعات نلاحظ التشابه الكبير بين وحداتها الذي يمكن أن يصل إلى درجة التجانس ، وفي بعضها نلاحظ الاختلاف الكبير ، وحيث أن العينة يجب أن تكون صورة صادقة من المجتمع الذي تسحب منه فمن البديهي ان حجم العينة بالنسبة للمجتمعات التي هي من النوع الأول يمكن أن يكون حجماً صغيراً بينما لا بد أن يكون حجماً كبيراً بالنسبة للمجتمعات التي هي من النوع الثاني ، وبمعنى آخر يتناسب حجم العينة تناسباً طردياً مع درجة الاختلاف بين وحدات المجتمع الذي تمثله . وبذلك لا بد أن نتعرف على درجة الاختلاف بين وحدات المجتمع حتى نستطيع أن نحدد حجم العينة التي يجب سحبها منه للتوصل إلى نتائج ذات دقة معينة . هذا المثال البسيط يوضح لنا أن سحب العينات لا يمكن أن يجري مستقلاً بذاته دون بعض المعرفة لخصائص المجتمعات التي نريد دراستها بالمعينة .

٢ - ان الهدف الأساسي عند سحب العينات هو التوصل إلى مجموعة من الوحدات ، التي يمكن أن تكون مع بعضها صورة صادقة من المجتمع الذي نرغب في دراسته حتى يمكن ان تكون اخطاء المعينة في النتائج اقل ما يمكن وبذلك تزداد اهميتها كمقاييس يمكن الاعتماد عليها في النواحي العملية المختلفة

(التخطيط مثلا) . لذلك يواجه الباحث عند سحب اية عينة مشكلة اساسية تتلخص في كيفية منع التحيز في الاختيار حتى يستطيع ان يطمئن الى العينة المسحوبة . لذلك يكون من المفيد ان تناقش العوامل المختلفة التي يمكن ان تؤدي الى التحيز في العينة حتى نستطيع ان نتفادها عند قيامنا بالعمل .

أ - استخدام اطار ^(١) غير شامل لوحدات المجتمع جميعاً ، ولا شك ان التحيز الناتج يكون خطيراً اذا كانت الوحدات التي لا يشملها الاطار من نوع معين له علاقة بموضوع البحث ، مثلاً امال الاسر التي تسكن في العشش في دراسة عن مستوى المعيشة بين اسر مدنية ما . ان عدم ادراج هذه الاسر ضمن الاطار لا يعطيها

(١) اطار المعاينة هو مجموع الوحدات التي نسحب منها العينة، أي مجتمع المعاينة، الوحدة في هذا الاطار هي وحدة المعاينة . ويمكن أن يكون مجتمع المعاينة هو نفسه المجتمع الاحصائي موضوع البحث أو مجتمع آخر يختلف عنه ، فمثلاً عند سحب عينة من الأسر يمكن أن يكون الاطار عبارة عن قائمة باسماء اسر المجتمع موضوع البحث وعناوين هذه الاسر، ويمكن أن يكون الاطار عبارة عن خريطة تظهر فيها بلوكات المساكن التي تسكنها هذه الاسر ، وكل بلوك منها يكون هو وحدة المعاينة ومجموع البلوكات هو مجتمع المعاينة . وبالطبع يسحب عدد من البلوكات كعينة ثم تجمع المعلومات عن الاسر التي تسكن في هذه البلوكات . وعندما نريد سحب عينة من الحيازات الزراعية يمكن أن يكون اطار المعاينة هو قائمة باسماء الحائزين وعناوينهم ويمكن أن يكون قائمة باسماء القرى التي نسحب بعضها كعينة ثم تجمع المعلومات من الحيازات في هذه القرى المسحوبة أو من عينة منها .

وبذلك نلاحظ ان اطار المعاينة يمكن أن يكون قائمة باسماء وحدات أو خريطة مقسمة الى مساحات معينة أو بلوكات معينة ، كما يمكن أن يكون صورة شمسية مأخوذة من الجو (في حالة العينات من الأراضي الزراعية بشرط أن تؤخذ الصور على ارتفاع ثابت دائماً ولذلك لا يمكن استعمال هذه الصور كاطر للمعاينة الا في الدول التي تكون فيها الأراضي الزراعية مستوية غير جبلية) . ومهما كان نوع الاطار يجب أن يشمل على جميع وحدات المجتمع موضوع البحث وعلى بعض الصفات الاساسية لمحدد الوحدات كعناوينها حتى يمكن التعرف عليها عند سحبها في العينة، ولكأي صفة اخرى تساعد في التعرف على درجة الاختلاف بين وحدات المجتمع حتى يمكن تحديد نوع العينة التي تناسب هذا المجتمع وعدد الوحدات التي يجب سحبها في العينة .

اية فرصة لان تمثل في العينة ، الامر الذي لا بد ان يؤدي إلى تحيز النتائج الى أعلى ، وتكون الدراسة بذلك دون أية قيمة علمية . ونلاحظ ان عدم شمول الاطار لا يؤدي فقط الى التحيز في سحب العينة وانما يؤدي كذلك الى الخطأ في تقدير النتائج النهائية . فاذا افترضنا أن مجتمع ما يتكون من ١٠٠٠ اسرة مثلاً في حين ان عدد الاسر في قائمة الاطار هي ٩٠٠ فقط (دون ان يعرف الباحث ذلك) وسحب الباحث ١٠٠ اسرة كعينة ، فانه بذلك سوف يقوم بتقدير صفات المجتمع موضوع البحث من العينة على أساس أن كسرهما $\frac{1}{9}$ بينما يكون الكسر الحقيقي للعينة $\frac{1}{10}$

وبذلك تكون التقديرات متحيزة إلى أسفل . كذلك يمكن أن يكون الاطار غير شامل بسبب قدمه مثل القوائم التي تؤخذ من الادارات الحكومية المختلفة عن المؤسسات الصناعية او التجارية أو غيرها من المؤسسات حيث تكون هذه القوائم في الغالب قديمة فلا تشمل على المؤسسات الجديدة التي انشأت بعد تاريخ اعداد هذه القوائم . كذلك يمكن ان يكون الاطار غير شامل عند استعمال دليل التلفون حيث لا يكون لدى جميع اسر المجتمع تليفونات ، الأمر الذي يجعل الدراسة متحيزة ضد هذه الاسر وفي صالح الاسر التي يكون لديها تليفونات .

ب - استخدام اطار به ازدواج أي تتكرر فيه أسماء بعض الوحدات وطبعاً يكون التحيز خطيراً اذا كانت هذه الوحدات التي تكررت اسماؤها تمثل نوعاً معيناً من الوحدات له علاقة بموضوع البحث . ويحدث هذا غالباً في القوائم التي تظهر أسماء المحلات التجارية حيث يكون لهذه المحلات فروع مختلفة في احياء مختلفة باسماء مختلفة ، وحيث أن العدادين الذين يقومون بحصر المحلات يوزعون على الأحياء المختلفة لاجراء عملية الحصر فاذا لم يكن لديهم التعليمات بسؤال أصحاب المحلات التجارية عما إذا كان لديهم فروع في أحياء مختلفة ، فان القوائم التي يعدونها سوف

يظهر فيها الازدراج . والتعيز الذي ينتج عن الازدواج يكون بسبب أن الوحدات التي تتكرر اسماءها تعطى أكثر من فرصة للدخول في العينة بينما لا يكون للوحدات الأخرى إلا فرصة واحدة فقط. كذلك يترتب على الازدواج التعيز في التقديرات النهائية ، حيث إذا افترضنا ان القائمة باسماء المحلات التجارية تحتوي على ١٠٠٠ وحدة منها ١٠٠ وحدة عبارة عن وحدات متكررة ، أي ان العدد الحقيقي للمحلات التجارية هو ٩٠٠ فقط (دون أن يعرف القائمون على البحث) وسحبنا من هذا الاطار ١٠٠ وحدة كعينة فاندنا في النهاية سوف نقوم بالتقدير على أساس ان كسر العينة = $\frac{1}{10}$ بينما

يكون الكسر الحقيقي = $\frac{1}{9}$ وبذلك تكون التقديرات متحيزة الى أعلى .

ج - سيطرة الانسان يحسمه أو بعقله على سحب وحدات العينة ، حيث أن الانسان مهما كان محايداً لا بد أن يتأثر بعوامل معينة عند سحب العينة وقد أجريت التجارب الكثيرة التي اثبتت هذه الحقيقة ، فاختيار الانسان لارقام يمكن أن يؤدي الى التعيز حيث انه قد يكره بعض الأرقام لدلالاتها السيئة بالنسبة له وقد يحب بعض الأرقام لدلالاتها الحسنة بالنسبة له وبذلك لا يمكن أن يكون الانسان عشوائياً صحيحاً في اختيار الأرقام اذا كان مسيطراً على عملية السحب بعقله الخاص . كذلك يكون سحب بطاقات من كيس به بطاقات متساوية الشكل واللون عملاً متحيزاً حيث أنه في الغالب تلتصق البطاقات ببعضها ، الأمر الذي يعطي البطاقات الخارجية فرصة الدخول في العينة ولا يعطي البطاقات المضغوطة بينها أية فرصة . وحتى إذا كان عدد البطاقات صغيراً يكون من الأفضل عدم اتباع هذه الطريقة نظراً لاحتمال التعيز مهما كان هذا الاحتمال صغيراً .

د - عدم اجابة بعض الوحدات التي سحبت في العينة أو رفضهم الإجابة

أربعة أو أي عدد آخر من الأرقام تستلزمه عملية سحب العينة العشوائية . ونلاحظ أن أي عدد نكوته من هذه الأرقام يكون عدداً عشوائياً لأن كل خانة منه عشوائية ووضعها بجانب الأخرى عشوائياً كذلك . وحتى لا يكون للشخص حرية الاختيار اذا تقرر قراءة هذه الأرقام العشوائية بطريقة هندسية معينة أي في أي اتجاه معين يجب عدم تغيير اتجاه القراءة كلما تراءى لنا ذلك . كذلك نلاحظ أنه اذا تكرر العدد عند القراءة نستبعده كما انه اذا قرأنا رقماً لا يدخل ضمن أرقام المجتمع موضوع البحث نستبعده أيضاً . وبذلك تكون خطوات العمل كالآتي :

١ - وضع وحدات المجتمع في قائمة مرتبة حسب الأحرف الهجائية ، واذا كان اطار المعاينة خريطة أو صورة شمسية لا يكون هناك داعي لهذه الخطوة .

٢ - اعطاء هذه الوحدات أرقاماً متسلسلة .

٣ - على اساس عدد الأرقام في أكبر عدد متسلسل نصل اليه نقرأ الأرقام العشوائية ، فاذا كان أكبر عدد مكون من رقمين نقرأ الأرقام العشوائية رقمين رقمين ... وإذا كان أكبر عدد مكون من ثلاث أرقام نقرأ الأرقام العشوائية ثلاثة أرقام . ويمكن أن نبدأ في القراءة من أي مكان في الجدول على أن نستمر في القراءة حسب نظام هندسي ثابت .

٤ - نستبعد الأعداد التي تصادفنا والتي تكون أكبر من أكبر رقم متسلسل لدينا .

٥ - نستبعد الأعداد التي تصادفنا والتي سبق أن ظهرت في سحب العينة .

٦ - اذا كانت الأرقام المتسلسلة في إطار العينة تصل الى ٢٠٠ مثلاً فاننا نقرأ الأعداد العشوائية من ثلاث خانات ولكننا بذلك سوف نقرأ أرقام كثيرة مكونة من ثلاث خانات ولكنها تزيد على ٢٠٠ فنستبعدها ، إلا ان ذلك سوف يضيع علينا وقتاً طويلاً . لذلك نقسم مجموعة الأعداد العشوائية

الى مجموعات من ١ الى ٢٠٠ ومن ٢٠١ الى ٤٠٠ ومن ٤٠١ الى ٦٠٠ ومن ٦٠١ الى ٨٠٠ ومن ٨٠١ الى ٩٩٩ ثم نترك المجموعة الاولى كما هي ثم نحول كل من المجموعات الباقية الى المجموعة الاولى وذلك بطرح ٢٠٠ من المجموعة الثانية و ٤٠٠ من المجموعة الثالثة و ٦٠٠ من المجموعة الرابعة و ٨٠٠ من المجموعة الخامسة ، وبذلك نكون قد حولنا جميع الأعداد العشوائية المكونة من ثلاث خانات والتي تزيد من ٢٠٠ إلى أعداد عشوائية من ١ الى ٢٠٠ ، فإذا قرأنا مثلاً ٢٣٥ يكون العدد العشوائي الذي نأخذه هو ٣٥ وإذا قرأنا العدد العشوائي ٧٢٨ يكون العدد العشوائي الذي نأخذه هو ١٢٨ . وبهذه الطريقة نكون قد أخذنا العينة من أعداد متتالية في الجدول ووفرنا الوقت اللازم للاستمرار في القراءة حتى نحصل على العينة المطلوبة .

٧ - كذلك إذا كان لدينا مساكن في شوارع مختلفة وأردنا أن نسحب منها عينة يكفي أن نعرف عدد المساكن في كل شارع حتى نسحب العينة التي نريدها . نفترض ان توزيع المساكن هو كالآتي :

الشارع	رقم	١	٧٥	مسكن
»	»	٢	١٢	»
»	»	٣	١٩٦	»
»	»	٤	٣٢	»
»	»	٥	٦٣	»
»	»	٦	١٤	»

تكون متجمعا صاعداً لعدد المساكن في هذه الشوارع كالآتي :

الشارع	رقم	١	٧٥	مسكن
الشارع الاول + الشارع	رقم	٢	٨٧	مسكن
الشارعين الاول + الشارع	رقم	٣	٢٨٣	مسكن

الثلاث الشوارع الاولى + الشارع رقم ٤	٣١٥	مسكن
الاربع الشوارع الاولى + الشارع رقم ٥	٣٧٨	مسكن
الخمس الشوارع الاولى + الشارع رقم ٦	٣٩٢	مسكن

وبذلك نعرف اننا اذا رقنا المساكن ترقياً متسلسلاً فسوف يكون آخر عدد هو ٣٩٢ وهو مكون من ثلاث خانات . نقرأ الاعداد العشوائية من ثلاث خانات ونقسمها الى مجموعات تبدأ المجموعة الاولى من ١ الى ٥٠٠ والثانية من ٥٠١ الى ٩٩٩ وتبقى المجموعة الاولى كما هي ونطرح ٥٠٠ من اعداد المجموعة الثانية . نفترض اننا قرأنا العدد العشوائي ٠٨٦ فان هذا العدد يقع في الشارع الثاني ويكون رقم المسكن في هذا الشارع هو رقم :

٨٦ - ٧٥ = ١١ . نفترض اننا قرأنا العدد العشوائي ٧٦٤ نطرح منه ٥٠٠ يكون العدد الباقي هو ٢٦٤ وهذا العدد يقع في الشارع الثالث ورقم المسكن في هذا الشارع هو ٢٦٤ - ٨٧ = ١٧٧ . وهكذا حتى تكون العينة التي نريدها .

٨- نفترض ان لدينا حقل مستطيل طوله ٣٠٠ متر وعرضه ٢٠٠ متر وأردنا ان نأخذ منه عينة من ثلاث قطع مساحة كل منها متر مربع . نقسم طول الحقل الى امتار تبدأ من الاعداد صفر حتى ٣٠٠ وكذلك نقسم عرض الحقل الى امتار تبدأ من العدد صفر حتى ٢٠٠ . نطرح من الطول متر واحد فيكون آخر عدد على الطول هو ٢٩٩ ، وكذلك نطرح من العرض متر واحد فيكون آخر عدد على العرض هو ١٩٩ . نقرأ الاعداد العشوائية من ثلاث ارقام فاذا ظهرت اكثر من ٣٠٠ نطرح هذا العدد من العدد الذي قرأناه حتى العدد ٥٩٩ واذا ظهر العدد من ٦٠٠ الى ٨٩٩ نطرح منه ٦٠٠ ويستبعد اي عدد يزيد على ٨٩٩ . وكذلك بالنسبة للعرض . ان قراءتنا الاولى تعطينا عدداً على الطول وقراءتنا الثانية تعطينا عدداً على العرض . عند تقاطع المحورين المقامين من هذين العددين نأخذ مربعاً مساحته متر مربع في اتجاه

الترقيم على كل من الطول والعرض .

ان العينة المسحوبة بهذه الطريقة تسمى بالعينة العشوائية وهي التي يتم سحبها على اساس تكافؤ فرص الاختيار لجميع وحدات مجتمع المعاينة - الا ان مثل هذه العينة لها عيوب يمكن ان تظهر بالنسبة لبعض المجتمعات . فالوحدات المأخوذة يمكن ان تظهر على الطبيعة متناثرة تناثراً كبيراً بحيث يصبح العمل في الميدان صعباً جداً وبحيث يكون الاشراف على العدادين اثناء عملهم امراً متعذراً . ولا شك ان هذا العيب يكون بالغ الاهمية اذا كان المجتمع موضوع البحث مترامي الاطراف بحيث يكون تناثر وحدات العينة في هذا المجتمع امراً له أهميته . الا ان العكس يمكن ان يحدث اذ قد تتركز وحدات العينة في قسم من المجتمع ويكون هذا التركيز امراً خطيراً حيث يكون عاملاً مؤدياً الى التحيز اذا كانت اقسام المجتمع تختلف عن بعضها اختلافاً جذرياً ، فاذا أردنا مثلاً اختيار عينة عشوائية من الاراضي الزراعية في لبنان ونحن نعلم ان اراضيها تنقسم الى اجزاء (اراضي الساحل و اراضي الجبل و اراضي البقاع) تختلف عن بعضها اختلافات كثيرة من حيث التربة و انواع المحاصيل التي تزرع في كل جزء منها فان هذه العينة يحتمل ان تتركز في ابي جزء من هذه الاجزاء ولا تكون بذلك صورة صادقة من هذا المجتمع بل انها بالعكس تكون متحيزة تحيزاً صارخاً . لذلك لا يمكن ان يطمئن صاحب العينة الى العينة العشوائية اذا كان المجتمع موضوع البحث من هذا النوع . كذلك عندما نريد سحب عينة عشوائية لا بد من اعداد اطار شامل لجميع وحدات المجتمع ، وهو امر قد يكون من المتعذر في بعض الاحيان نظراً لان اعداد مثل هذا الاطار يحتاج الى حصر وحدات المجتمع حصر شاملاً وذلك يحتاج الى تكاليف باهظة قد لا يكون الجهاز القائم بالعمل غير مستعد لتحملها . ومن ناحية اخرى رأينا ان سحب العينة باستخدام الارقام العشوائية فيه ضياع لوقت الساحب في قراءة الأرقام العشوائية خاصة إذا كثرت وحدات المجتمع وكانت العينة

التي نريد سحبها ذات حجم كبير . لذلك قد يكون من الافضل التفكير في نوع آخر من انواع العينات التي فكر فيها علماء الاحصاء لمواجهة مثل هذه العيوب في العينة العشوائية البسيطة .

وباستخدام العينة الطبقية يمكن معالجة الاختلاف الكبير بين وحدات المجتمع . ويكون ذلك بتقسيم المجتمع إلى طبقات متجانسة ، ففي مثالنا عن الاراضي الزراعية في لبنان تقسم الاراضي الى ثلاث أقسام - أراضي الساحل وأراضي الجبل وأراضي البقاع . وفي بعض الحالات يكون انقسام المجتمع إلى طبقات أمر موجود فعلاً في الطبيعة كما هو الحال بالنسبة للأراضي الزراعية في لبنان ، إلا انه في حالات أخرى يتحتم على الباحث نفسه اجراء هذا التقسيم . وفي مثل هذه الحالات يجب أن يكون لديه صفة معينة أو مقياس معين يمكن على أساسه اجراء التقسيم حيث ان الهدف من التقسيم هو تجميع كل مجموعة من الوحدات المتجانسة في قسم أي طبقة قائمة بذاتها . فاذا كنا ندرس اجور العمال يكون من الأفضل تقسيمهم حسب المهن التي يقومون بها ، وإذا كنا ندرس الانتاج الصناعي يكون من الأفضل تقسيم المؤسسات الصناعية حسب نوع نشاطها وحجم المؤسسات التي تعمل في كل نشاط . بعد تقسيم المجتمع إلى طبقات تؤخذ من كل طبقة عينة اما بالسحب العشوائي أي باستخدام الاعداد العشوائية أو بالسحب المنتظم الذي سنشير اليه فيما بعد . ثم تدرس كل عينة على حدة ثم تعمم النتائج على المجتمع كله مع ترجيح نتائج كل طبقة بحسب نسبة العينة المأخوذة منها .

بهذه الطريقة نحقق مزايا كثيرة ، فأولاً نؤكد من بادىء الأمر عدم تحيز العينة حيث انها سوف تقسم على طبقات المجتمع المختلفة فلا خوف من تركزها في طبقة دون أخرى كما هو الحال بالنسبة للعينة العشوائية البسيطة ولا شك أن ذلك يؤدي بطريق غير مباشر الى زيادة دقة النتائج . وثانياً يمكننا أن نحصل على صفات كل جزء من المجتمع على حدة بجانب معرفة صفات المجتمع

ككل غير مجزا . ففي مثالنا عن الاراضي الزراعية نستطيع أن نتعرف على صفات الاراضي الزراعية في كل من الساحل والجبل والبقاع بجانب معرفة صفات الاراضي الزراعية في مجموعها بعد ترجيح النتائج الخاصة بكل منطقة على حدة .

وإذا كان لدينا إطار عام بوحدات المجتمع جميعاً مابين به الصفة التي على أساسها سوف نسحب عينة طبقية . فكل ما نريد ان نعرفه هو عدد الوحدات التي نريد سحبها في كل طبقة وبذلك نعد إطارات فرعية لكل طبقة على حدة ، ثم بالقراءة من الاعداد العشوائية إذا ظهرت وحدة ما نعرف صفتها ، وعلى هذا الاساس نضعها في إطار الطبقة الخاصة بها ونستمر في العمل بهذه الطريقة حتى تكون الوحدات المطلوبة في كل طبقة على حدة .

وكما سبق أن ذكرنا يتحدد حجم العينة تبعاً لدرجة الاختلاف بين وحدات المجتمع ، ولهذا اذا كان المجتمع شديد الاختلاف فان العينة العشوائية البسيطة التي يجب سحبها من هذا المجتمع لا بد وأن تكون كبيرة جداً حتى تعطينا درجة دقة معينة . إلا انه بتقسيم هذا المجتمع الى طبقات بحيث تشتمل كل طبقة على الوحدات المتجانسة فان العينة التي يجب سحبها من كل طبقة سوف تكون صغيرة بحيث يكون مجموع الوحدات في العينات التي تؤخذ من الطبقات المختلفة أقل بكثير من حجم العينة العشوائية البسيطة ، وهذه ميزة هامة للعينة الطبقية حيث يمكن على اساسها انقاص تكاليف الدراسة الى أدنى حد ممكن . على انه يجب أن نلاحظ ان هذه الميزة لا تتحقق الا اذا كان تقسيم المجتمع الى طبقات يؤدي الى نقص الاختلاف في كل طبقة على حدة نقصاً كبيراً ؛ أما إذا كان النقص في الاختلاف سوف يكون بسيطاً فان حجم العينة الطبقية يحتمل ان يزيد عن حجم العينة العشوائية البسيطة .

وباستخدام العينات المنتظمة نستطيع أن نختصر كثيراً من الوقت الذي يضيع في السحب باستخدام الأعداد العشوائية . والصفة الاساسية لهذه العينات

هي تساوي المسافات في الاطار بين الوحدات المتتالية التي تؤخذ في العينة ولاستخدام هذه الطريقة في سحب العينات يجب أن نعرف عدد وحدات المجتمع ، وليكن ١٠٠٠ مثلا فاذا أردنا أن نسحب ١٠٠ وحدة كعينة، فإننا

نقسم حجم المجتمع على حجم العينة $\frac{1000}{100} = 10$ وبذلك نحصل على ما نسميه

مسافة الانتظام وهي ١٠ في هذا المثال . ويعني ذلك انه لا بد أن نأخذ وحدة من كل عشرة وحدات من المجتمع . وبذلك نبدأ بقراءة عدد عشوائي حتى العدد عشرة فاذا فرضنا ان هذا العدد هو ٧ يكون العدد التالي لوحدة العينة ١٧ والعدد الثالث ٢٧ وهكذا .

ولست ميزة سهولة السحب هي الميزة الوحيدة للعينة المنتظمة، ففي كثير من الأحيان تكون أصدق تمثيلا للمجتمع أي أقل تأثراً بخطأ الصدفة من العينات العشوائية البسيطة فأخذ عينة منتظمة من المنازل في إحدى المدن يضمن لنا تمثيل العينة لمنازل الأحياء المختلفة وبنسبة تساوي تقريبا نسبة عدد المنازل في كل حي . وبهذا تكون العينة المنتظمة اذق تمثيلا في هذه الحالة من العينة العشوائية فيما يختص بالنواحي الاجتماعية والاقتصادية والعمرائية المرتبطة بتوزيع المنازل على الأحياء المختلفة في المدينة .

على انه يجب أن نلاحظ عند سحب العينات المنتظمة عدم انتظام الوحدات في اطار المعاينة انتظاماً له علاقة بموضوع البحث ، وقد سبق أن أشرنا الى ذلك فيما يختص بترتيب العمال في قائمة حيث يمكن أن يكونوا منظمين في مجموعات كل مجموعة منها لها رئيس الذي يظهر اسمه بشكل منتظم في أول كل مجموعة . كذلك لا يجب أن نظن ان العينة المنتظمة ليست الا عينة عشوائية حيث نسحب الوحدة الاولى عشوائياً أي باستخدام الاعداد العشوائية . ولكن يجب الا ننسى أن ليس هناك في العينة تعشوائية ما يمنع من ظهور أي وحدة منها كان ترتيبها في الاطار ، بينما في العينة المنتظمة اذا تحددت الوحدة

الاولى فان باقي الوحدات في العينة تتحدد على أساس مسافة الانتظام التي تتبعها. كذلك نلاحظ أن عدد العينات التي يمكن سحبها باستخدام السحب العشوائي يكون عدد كبير جداً يصل الى ما لا نهاية ، أما بالسحب المنتظم فان عدد العينات التي يمكن سحبها عدد محدد ، فاذا كانت مسافة الانتظام ١٠ كان عدد العينات التي يمكن سحبها عشرة عينات فقط واذا كانت مسافة الانتظام ٢٠ فان عدد العينات التي يمكن سحبها عشرين عينة فقط . ولذلك نقابل في العينات المنتظمة صعوبات تحليلية حيث أن دراسة خطأ الصدفة على النتائج أصعب عادة في هذه الحالة منه في حالة العينات العشوائية .

وباستخدام العينات على مراحل نستطيع ان نتغلب على صعوبة اعداد اطار للمجتمع كله خاصة اذا كان هذا المجتمع واسع جداً بسبب الصعوبات المالية والادارية التي نواجهها في هذه الحالة . ففي بحث عن الحيازات الزراعية قد نجد انه من العبث أن نفكر في أخذ عينة عشوائية أو منتظمة في الدولة بأكملها . وفي مثل هذه الحالة يتعين علينا تجميع العينة في مناطق محدودة حتى نستطيع أن نتغلب على الصعوبات المالية والادارية، ويكون ذلك بأخذ عينة على مراحل وفي كل مرحلة نصغر وحدة العينة عن المرحلة السابقة لها حتى ننتهي بوحدة البحث في المرحلة الأخيرة . ففي المرحلة الأولى نأخذ عينة من الأقضية ومن عينة الأقضية نأخذ عينة من القرى ومن القرى نأخذ عينة من الحيازات الزراعية التي نريد دراستها .

٦ - تقدير الميزانية الزمنية والمالية والبشرية الضرورية للبحث -

إن عمليات بحوث الميدان تكون من التداخل والتشابك بحيث لا بد من إعداد جدولاً زمنياً تبين فيه العمليات المختلفة التي يتكون منها العمل وتوضع أمام كل عملية خط يبين متى يجب أن تبدأ ومتى يجب أن تنتهي وبذلك يبين المدة اللازمة لانهاؤها . والهدف الأساسي من هذا الجدول أن يطمئن الباحث في أي وقت يشاء الى ما إذا كان يستمكن من الانتهاء من بحثه ونشر نتائجه

في الموعد المحدد ، ولكي يعرف مواطن التأخير فيعمل على تعجيلها حتى لا تتراكم التأخيرات وتؤدي في النهاية إلى عجز الباحث عن المحافظة على المواعيد المقررة . ولا بد عند اعداد هذا الجدول أن نراعي كل العوامل التي يمكن ان تؤثر في سير العمل كالمرض والفصل والاجهاد والعطلات وظروف الطقس وما إلى ذلك من عوامل يمكن ان يكون لها تأثير مها بلغ الباحث من الدقة في التصميم والاحكام .

كذلك يجب أن يلتزم الباحث ببرنامج انفاقي حتى يطمئن الى امكان اتمام البحث والسير فيه حتى نهايته . ولا يعني ذلك تخصيص مبلغ معين للبحث وانما المقصود هو اعداد ميزانية مفصلة تظهر المبالغ المقدرة لكل عملية من العمليات التي يتكون منها البحث . وفي الغالب تعد ميزانيات البحوث على أساس تقسيم العمل الى ثلاث مراحل - مرحلة الاعداد للنزول الى الميدان وتتضمن وضع خطة البحث وتصميم الاستمارات اللازمة وصياغة التعليمات والارشادات وتجنيد وتدريب العدادين والمشرفين على الأعمال التي سوف توكل اليهم والقيام بالبحث التجريبي ، ومرحلة النزول الى الميدان لجمع المعلومات ومراجعتها ، ومرحلة التبويب وتشمل عمليات الترميز والتثقيب والفرز والتبويب وتحليل البيانات وكتابة التقرير ونشره ان أعداد ميزانية مفصلة لهذه العمليات المختلفة يساعد في التوصل الى تقدير للرقم الاجمالي للتكاليف دون أن يكون ذلك مبنياً على الحدس والتخمين . كذلك تساعد الميزانية التفصيلية في مراقبة الانفاق على العمليات المختلفة للتأكد من عدم التلاعب في ذلك . واعداد الميزانية التفصيلية لا يعني تقدير مبالغ جامدة لا يمكن تغييرها فيما بعد ، حيث يمكن أن تستجد ظروف مثل ارتفاع أسعار الورق تؤدي الى ضرورة التغيير . لذلك بعد أن ينتهي العمل لا بد من أعداد حساب ختامي يظهر ما أنفق فعلاً على كل عملية من العمليات حتى يكون هذا الحساب أساساً سليماً نرتكز عليه عند اعداد ميزانيات البحوث في المستقبل .

وبناء على القرارات السابقة فيما يختص بالمعلومات التي يرغب الباحث في

جمعها ، ونطاق البحث وشموله ، ومجال جمع المعلومات ، والجدول الزمني يستطيع الباحث أن يقدر الميزانية البشرية أي عدد الموظفين الضروريين للقيام بكل عملية يتطلبها البحث .

ثانياً - تصميم الاستمارة والتعليقات الخاصة بها :

عند تصميم الاستمارة الخاصة بموضوع البحث يجب ملاحظة الآتي :

١- لا يجب أن تشتمل الاستمارة إلا على الأسئلة التي لها علاقة مباشرة بموضوع الدراسة ، ولذلك يجب الموازنة بين الفائدة التي تعود من اضافة سؤال ما والزيادة في التكاليف والوقت والجهود التي تنشأ عن ذلك ، بحيث اذا تبين ان الفائدة منه لا تتناسب وزيادة التكاليف وجب اهماله وعدم ادراجه في الاستمارة . ولكي يتأكد الباحث أن الاستمارة ستكون شاملة لجميع البيانات اللازمة لتحقيق أغراض الدراسة يجب اعداد الجداول التخيلية التي سوف تصنف وتبوب فيها المعلومات بعد جمعها . وعند التفكير في أسئلة الاستمارة يجب أن تكون هذه الجداول في مخيلة المصمم دائماً بحيث يستطيع أن يحدد كيف ستظهر اجابة كل سؤال كبيان مصنف في واحد من هذه الجداول أو في بعضها . وبهذه الطريقة في التصميم يمكن أن يتقرر اهمال بعض الأسئلة أو اعادة صياغة بعضها أو التفكير في أسئلة اخرى .

٢- لا يجب أن تحتوي الاستمارة إلا على الأسئلة التي يكون جميع المستجوبين أو غالبيتهم على الأقل قادرين على الاجابة عليها ، حيث يجب أن يؤخذ عامل الزمن في الاعتبار اذا يمكن أن تكون الاجابة قد توفرت لدى المستجوب في وقت ماض الا أنه نسيها بفعل مرور الزمن فلا يصح مثلاً أن نسأل عن دخل المستجوب قبل أعوام مضت أو عدد ساعات العمل التي عملها في الماضي .

٣- يجب تجنب الأسئلة التي يكون فيها حرج للمستجوب نظر لعلاقتها بأحواله الشخصية ، وبذلك يجب الابتعاد عن الأسئلة التي يعتبرها المستجوب

تدخل في حياته الخاصة أو لها مساس بتقاليده وقيمة الأخلاقية . وإذا اضطرب الباحث الى مثل هذا النوع من الأسئلة الشخصية يحسن توظيف عدادين ليسوا من أهل المنطقة التي يعملون فيها، كما يحسن إدراج هذه الأسئلة بين أسئلة أخرى كثيرة ليس لها الصفة الشخصية .

٤ - يجب تجنب الأسئلة التي تستدعي الإجابة عليها اجراء عمليات حسابية من قبل المستجوبين اذ غالباً لا يعطون اجابات دقيقة عليها . فالسؤال عما تنفقه الاسرة شهرياً على الملابس والمصروفات المدرسية يستدعي اجراء عمليات حسابية نظراً لان هذه النفقات تكون في الغالب فصلية وليست شهرية .

٥ - يجب تجنب الأسئلة الایجابية أي تلك التي توحى باجابة معينة يريد الباحث اثبات صحتها ، ومن الأسئلة على ذلك - هل تغيرت بسبب المرض ؟ هل تأخرت بسبب سوء المواصلات ؟ هل سرقت لشعورك بالجوع ؟ والاجابة على مثل هذه الأسئلة يكون في الغالب بالايجاب .

٦ - يجب اختيار الالفاظ البسيطة السهلة ، كما يحسن اختيار الشائع منها بين المستجوبين، لذلك يجب أن يؤخذ المستوى التعليمي للمستجوبين في الاعتبار عند وضع الأسئلة . ان الالفاظ مثل الرسوم الجهركية ، الضرائب غير المباشرة ، اليساريين ، الحالة الزوجية ، النشاط الاقتصادي ، البطالة ، ... الخ قد تبدو واضحة في ذهن الباحث ولكنها بالتأكيد تكون غير مفهومة بتاتا من بعض المستجوبين وغير واضحة المعنى بالنسبة للبعض الآخر . وإذا اضطرب مصمم الاستمارة الى استخدام هذا النوع من الالفاظ يجب أن يشرح معناها في التعليمات المفسرة للاستمارة .

٧ - لا يجب أن تكون الأسئلة طويلة، إذ قد تكون الالفاظ التي يتكون منها السؤال بسيطة للغاية الا أن المستجوب لا يدرك المقصود منها نظراً لطول السؤال وشموله لمقاطع كثيرة . ان المستجوبين غالباً يطالبون العدادين بإعادة

قراءة مثل هذه الأسئلة مراراً ، الأمر الذي يؤدي الى اضاعة الوقت ، وفي بعض الأحيان يكون طول السؤال يسبب ادماج سؤالين أو أكثر في سؤال واحد ، الأمر الذي يؤدي الى ارتباك المستجوب وارتباك العمل في المكتب عند اجراء التبويب ، فسؤال مثل هل انت عضو في أحد النوادي الرياضية وما هي الرياضة التي تمارسها وأياها المفضلة عندك ؟ هذا السؤال يجب تقسيمه الى ثلاث أسئلة فيصبح كل منها قصيراً وبسيطاً وواضحاً ، الأمر الذي يساعد العداد على القاء كل سؤال منها والمستجوب على فهمها بسرعة والاجابة عليها بدقة والمكتب على اجراء التبويب بسهولة ويسر .

٨ - يجب تجنب الأسئلة النفسية ، مثلاً هل تذهب الى السينما كثيراً ؟ سؤالي يترتب عليه اجابات بمعايير مختلفة حيث أن ما يعتبره البعض كثيراً يعتبره البعض الآخر قليلاً .

٩ - يجب تجنب الألفاظ المثيرة مثل الرأسمالية والشيوعية والمحافظين والأحرار واليساريين ورجال الأعمال والطبقة العاملة ... الخ ، حيث يكون لمثل هذه الألفاظ وقع مختلف على المستجوبين ، الامر الذي يوجد تحيزاً في اجاباتهم . كذلك لا يجب استخدام الاسماء الضخمة مثل هل توافق على رأي ... ؟ حيث أن عاطفة المستجوبين نحو هذه الاسماء تختلف ، الامر الذي يوجد تحيزاً في اجاباتهم . كذلك يجب الابتعاد عن الألفاظ الایحائية مثل كلمة « معقولة » حيث انه اذا استخدمت هذه الكلمة في سؤال مثل هل توافق على زيادة معقولة في الأسعار بأمل أن يؤدي ذلك الى الانتعاش ؟ على اساس هذا السؤال سوف يجيب نسبة كبيرة من المستجوبين بنعم وتقل هذه النسبة كثيراً اذا حذفنا كلمة معقولة من السؤال .

١٠ - يجب الحذر من استخدام الجمل التي يمكن أن يكون لها ارتباط بضمير المستجوب أو شعوره بالكرامة والفضيلة حيث ان غالبية الناس يرغبون في الشعور بانهم أذكياء ، منطقيين ، كرام ، صالحين ، محبين للغير ، وهم لذلك يميلون الى الاجابة على اساس ما يجب أن يكون سلوكهم أو رأيهم بالنسبة لوضع معين وليس على اساس سلوكهم أو رأيهم الواقعي .

١١ - كذلك يجب تجنب الأسئلة التي تدفع المستجوب الى الادعاء ، مثلاً السؤال هل قرأت كتاب كذا ؟ يمكن الاستعاضة عنه بالسؤال هل تنوي قراءة كتاب كذا ؟

١٢ - يحسن أن تتضمن الاستمارة بضع أسئلة يكون الغرض منها تدقيق اجابات الأسئلة الأخرى حيث يمكن السؤال عن نفس الشيء بطريقتين مختلفتين ووضع السؤالين في مكانين مختلفين من الاستمارة حتى يساعدنا ذلك في تدقيق التماسك الداخلي للإجابات ... فمثلاً يمكن التأكد من صحة الاجابة على سؤال عن العمر اذا وضع سؤال آخر عن العمر عند الزواج ومدة الحياة الزوجية او عن تاريخ الاقتراح للجنسية . وفي استمارة ميزانية الأسرة قد لا يكون من أغراض الدراسة جمع معلومات عن الدخل والاقتراض والادخار حيث يكون الغرض معرفة نفقات الأسرة فقط ومع ذلك يحسن إضافة أسئلة عن هذه العناصر الثلاث حتى تكتمل الصورة التي نأخذها عن الأسرة فيمكن بذلك تدقيق اجاباتها عن الانفاق .

١٣ - يحسن تحليل السؤال الى عناصره المختلفة ، مثلاً إذا كنا نسأل عند تفضيل المستجوب لنوع معين من السيارات يجب أن نتذكر ان هناك عوامل كثيرة للمفاضلة كالتكلفة والاداء والحجم والمظهر ومدى توفرها ، وكل من هذه العوامل له أيضاً جزئيات ، فالتكلفة مثلاً تنقسم الى قسمين قسم يتمثل في ثمن شراء السيارة وقسم آخر يتمثل في النفقات الجارية لاستخدامها . وبذلك يكون إجمال الاسئلة عن المفاضلة في سؤال واحد مؤدياً في الغالب الى إجابات مضللة .

١٤ - يجب صياغة الأسئلة بحيث لا تستدعي الاجابة عليها الا أقل ما يمكن من الكتابة ، ومعنى آخر ان تكون الاجابة عليها قابلة للتبويب في الجداول المعدة بكل سهولة ويسر ، ولذلك يحسن صياغة الاسئلة بحيث تكون الاجابه عليها باحدى الطرق الآتية :

١ - كتابة نعم أو لا . - ٢ - كتابة رقم معين . - ٣ - كتابة علامة معينة في مربع صغير خاص بإجابة معينة . - ٤ - وضع خط تحت الإجابة الصحيحة . - ٥ - إحاطة الإجابة الصحيحة بدائرة . - ٦ - كتابة رمز معين . - ٧ - كتابة كلمة أو جملة بسيطة . . على أن صياغة الأسئلة بحيث تكون الإجابة عليها بإحدى الطرق الواردة في الفقرات من ٣ الى ٥ تحتاج الى حصر جميع الاجابات المحتملة على كل سؤال بحيث لا يمكن أن تأتي إجابة أي من المستجوبين مختلفة عن الإجابات التي تقرر طبعها في الاستمارة . لهذا يجب التفكير في كل الاجابات المحتملة لكل سؤال دون أن ننسى احتمال الاجابة بـ « لا أعرف » أو « أشك » أو « لا أذكر » أو « لا رأي لي » أو « لا أستطيع الاختيار » . أما إذا رفض المستجوب الإجابة فيسجل العداد ذلك على السؤال دون الحاجة الى طبع هذه الإجابة المحتملة حيث ان طبعها في الاستمارة يشجع العداد على عدم بذل أي مجهود للحصول على إجابة معينة .

هذه بعض القواعد والشروط التي يجب ملاحظتها عند تصميم استمارة البحث ، واحب أن اشير إلى أن ادراك هذه القواعد لا يعني اننا نستطيع أن نصمم استمارة ما تصميا جيدا حيث أن هذا العمل يحتاج إلى خبرة طويلة ومعرفة صحيحة بعلم النفس وعلم الاجتماع . إلا أن ذلك لا يجب أن يخيفنا كثيراً حيث أن كثيراً من الاستمارات الاحصائية لا تواجه في تصميمها صعوبات كثيرة وهي تلك التي ترتبط بالبحوث الموضوعية مثل البحوث الخاصة بالصناعة أو الزراعة أو التجارة ... الخ حيث لا يكون لرأي المستجوب أي اعتبار في الاجابات التي نريدها ، إذ تتطلب الأسئلة في هذه البحوث الإجابة في شكل أرقام مقيدة في سجلات المستجوبين ، وحتى إذا لم تكن مقيدة لا يكون هناك أية صعوبة في الحصول عليها إذا أمكن اقناع المستجوبين بسرية المعلومات التي يعطونها وبأنها لن تستخدم ضدهم بأي حال من الأحوال وفقاً لقانون الاحصاء الذي سبق الإشارة اليه .

بعد صياغة أسئلة الاستشارة تكون الخطوة التالية هي تجميعها في مجموعات من حيث الترابط فيما بينها ، فإذا كنا بصدد بحث عن الصناعة مثلا ، يجب أن تكون الأسئلة الخاصة بعدد المشتغلين في المؤسسة منفصلة عن الأسئلة الخاصة بالاجور النقدية والعينية التي دفعت لهم وهذه منفصلة عن الأسئلة الخاصة بالانتاج وهكذا . وفي داخل كل مجموعة يجب ان نلاحظ التسلسل المنطقي للأسئلة حتى لا تنقطع سلسلة أفكار المستجوب اثناء اجابته . فلا يجب مثلا السؤال عن المستوى التعليمي للشخص إلا بعد السؤال عما إذا كان متعلما او اميا ، كما لا يجب السؤال عن مهنته إلا بعد السؤال عما إذا كان مشتغلا أو متعطلا . كذلك يجب ترك فراغ مناسب لكتابة الاجابة عن كل سؤال على حدة وفراغ آخر لكتابة الرموز اذا كان التبويب للاجابات سوف يجري آليا .

ولا يكتمل تصميم الاستشارة إلا باعداد التعليمات التي تفسر معاني الكلمات المختلفة التي استخدمت في الأسئلة والتي يمكن ان يساء فهمها من قبل المستجوبين ، ومعنى كل سؤال على حدة حتى يمكن أن يفهمه المستجوب نفس الفهم الذي يفهمه الباحث ، والتي تشرح كيفية الاجابة على كل سؤال ومتى تكون الاجابة صحيحة ومتى تكون خاطئة ، ثم كيفية كتابة الاجابة في الفراغ المخصص لها . فإذا كانت الاستشارة التي سوف نستخدمها هي صحيفة استبيان ترسل بالبريد للمستجوبين يحسن طبع التعليمات الخاصة بكل سؤال اسفله مباشرة ولكن باحرف أصغر وفي جمل مركزة قصيرة وذلك بجانب أعداد خطاب يرفق مع الاستشارة لتوضيح الغرض من البحث ومصرية البيانات وعدم استخدامها ضد صالح المستجوب ورجاء ملأ الاستشارة واعادتها إلى الباحث في تاريخ أقصاه كذا . أما إذا كانت الاستشارة التي سوف نستخدمها كشف البحث ، ففي هذه الحالة تطبع جميع التعليمات سواء الخاصة بالبحث في مجموعة من ناحية الغرض منه والسرية وسواء الخاصة بتفسير وتوضيح الاستشارة أو الخاصة بمقابلة المستجوب وكذلك الخاصة

بالنواحي الادارية لعمل العداد في الميدان في كتيب مستقل يوزع على العدادين
ليسترشدوا به عند قيامهم بعملهم

ثالثاً : تجنيد وتدريب العدادين :

إذا تقرر استخدام كشف البحث يكون الجهاز القائم بالعمل في حاجة إلى عدادين
يمكن أن يزيد عددهم إلى الآلاف في الدراسات بالعد الشامل ، الأمر الذي
يدعو إلى تجنيدهم من الطلبة والمدرسين وقت العطلات المدرسية ، ويمكن أن
يقل عددهم كثيراً في الدراسات بالمعينة وبذلك يمكن الاعتماد على موظفي
الجهاز الاحصائي الذين لديهم خبرة سابقة بهذا العمل .

ومهما كان عدد العدادين الذين يكون البحث في حاجة اليهم ، فان نجاح
البحث يتوقف أولاً وقبل كل شيء على استخدام العدادين الصالحين للعمل
الذي سوف يوكل اليهم . ولهذا يجب ملاحظة توفر الشروط الآتية في
توظيف العدادين :

- ١ - القدرة على التحدث بطلاقة مع طوائف المجتمع المختلفة .
- ٢ -- القدرة على جذب انتباه الناس بسرعة وبطريقة سليمة .
- ٣ - الملاحظة السريعة وادراك التفاصيل .
- ٤ - الاصرار على النجاح في العمل واثمائه بشكل كامل دون أي نقص
- ٥ - حب الناس ومعاملتهم بالعطف والطيبة .
- ٦ - الضمير الحي والأمانة وتحمل المسؤولية .
- ٧ - الافق الواسع دون الظهور بمظهر الشديد الذكاء .
- ٨ - الذاكرة الجيدة .

- ٩ - حب البحث - العقلية الباحثة .
- ١٠ - القدرة على استيعاب تعليقات كثيرة واتباعها بدقة .
- ١١ - القدرة على تلخيص ما يسمع من اجابات وتسجيلها تسجيلاً موضوعياً .
- ١٢ - تجنب التحيز الشخصي .
- ١٣ - الصحة الجيدة والطاقة الجسدية الفائقة .
- ١٤ - المظهر والسلوك الحسن .
- ١٥ - الخط الواضح .
- ١٦ - أن يكون حراً من أية مسئوليات أخرى بحيث يمكنه السفر إذا استدعى الأمر وبحيث يكون راغباً وقادراً على العمل في المساء وفي أيام الاجازات .
- ١٧ - أن يكون حائزاً على مؤهل علمي يتفق ومستوى الدراسة .
- ١٨ - يتوقف جنس العمداد على نوع الدراسة ، ففي الدراسات التي تتعلق بريات المنازل يحسن أن يكون العمدادين من الاثا بينا يحسن توظيف عدادين من المذكور في الدراسات التي تتعلق بغير ربات المنازل .
- ١٩ - من الأفضل أن يكون عمر العمداد بين ٢٥ و ٤٠ سنة حيث أن صغار السن ينقصهم القدرة على الاتصال بالناس بنجاح بينما لا يميل كبار السن إلى اتباع التعليمات بدقة بالاضافة إلى عدم ميلهم إلى الالحاح اللازم لحفز المستجوبين على قبول مقابلة العمداد والاستماع اليه واعطائه الاجابات المطلوبة .
- ٢٠ - في الدراسات الواسعة النطاق والتي تشمل أنحاء الدولة المختلفة تنشأ مشكلة تجنيد العمدادين محلياً (من كل منطقة للعمل فيها) أو مركزياً (توزيع العمدادين على المناطق المختلفة دون التقيد بمنطقة العمداد نفسه) .

ولكل من الطريقتين مزاياها وعيوبها ، فالتجنييد المحلي يساعد على ضمان معرفة العدد للجهة التي يعمل فيها ولعادات أهلها وتقاليدهم ، إلا أن المستجوبين يميلون غالباً إلى عدم اعطاء معلومات عن أنفسهم لاشخاص من منطقتهم خوفاً من استخدام هذه المعلومات مادة للحديث في المقامي والمجالس العامة ، ويظهر هذا الاتجاه بشكل واضح في الدراسات التي يسأل فيها عن معلومات يعتبرها المستجوبون أسراراً بالنسبة لهم . أما التجنييد المركزي فلا يحقق مزايا التجنييد المحلي التي سبق الإشارة إليها إلى ان المستجوبين لن يتخوفوا من افشاء العدادين لأسرارهم ، الأمر الذي يشجعهم على التعاون معهم . والمفاضلة بين الطريقتين تتوقف إلى حد كبير على نوع الدراسة فإذا لم تشمل الدراسة على أسئلة عن معلومات سرية بالنسبة للمستجوبين يكون من الأفضل اتباع الطريقة الأولى مثل تعداد السكان ، أما إذا اشتملت على مثل هذه المعلومات يحسن اتباع الطريقة الثانية مثل التعداد الزراعي .

على أن التقيد بالشروط السابقة يتوقف إلى حد كبير على عدد العدادين اللازم للعمل في الميدان ، حيث انه اذا كان العدد كبيراً فان الهيئة المشرفة على الدراسة سوف تجد نفسها مضطرة إلى التساهل في تطبيق هذه الشروط حتى تضمن حصولها على العدد المطلوب خاصة في الدول التي تجري فيها مثل هذه الدراسات لأول مرة . أما في الدراسات بالمعاينة فان الهيئة المشرفة على البحث يجب أن تتقيد إلى حد كبير بالشروط السابقة لتضمن نجاح الدراسة حيث أن تساهلها سوف يزيد من نسبة الأخطاء العامة التي يمكن أن تظهر بالإضافة إلى أخطاء المعاينة نفسها .

ويتوقف عدد العدادين الذين يحتاجهم البحث على العوامل الآتية . -

- ١ - عدد وحدات الدراسة ٢ - التوزيع الجغرافي لهذه الوحدات .
- ٣ - حالة المواصلات في الدولة . ٤ - طول الاستمارة أو قصرها . ٥ - نوع المعلومات المطلوب استعمالها (كشف بحث أو صحيفة استبيان) ، حيث في

حالة استخدام صحيفة اسنبيان يكون الجهاز في حاجة كذلك الى عدادين للملاحقة المستجوبين الذين لا يعيدون استماراتهم الى المكتب في الموعد المحدد لذلك .

واذا ما تم تجنيد العدادين الذين يحتاجهم البحث يبدأ تدريبهم ، ويتوقف نجاح برنامج التدريب على الذين سوف يوكل اليهم تدريب العدادين والاشراف عليهم في الميدان . لهذا يجب عند توظيف المشرفين التقيد بالشروط السابقة فيما يتعلق بالعدادين دون أي تساهل بالاضافة الى ضرورة أن يكون المشرفين من مستوى علمي أعلى ولديهم قدرة أكبر على الاستيعاب والبحث والملاحظة ومواجهة المشاكل على أن الصفة الأساسية التي يجب أخذها بعين الاعتبار عند توظيف المشرفين هي القدرة على قيادة الناس وتوجيههم وكسب احترامهم دون عنف بجانب القدرة على توضيح الامور وشرحها .

ويمكن أن يتخذ برنامج التدريب الخطوات التالية :

١ - اعداد المشرفين والقائمين على تدريب العدادين .

٢ - توزيع العدادين في حلقات للدرس (محلياً أو مركزياً) .

٣ - اعداد جميع المطبوعات اللازمة للتدريب وطبعها حتى يمكن توزيعها على العدادين وقت بدء التدريب . ويجب أن يتوفر لدى كل عداد نسخة من المحاضرات التي سوف تلقى اثناء التدريب ونسخة من كتيب التعليمات . أما غير ذلك من المطبوعات الادارية مثل كشوفات حصر وحدات الاستبيان في مناطق العد المختلفة ونماذج تقارير العمل اليومي فتوزع على العدادين عندما يتقرر قيامهم بالعمل . وقبل هذا أو ذاك يجب أن يكون لدى العداد بطاقة هوية تثبت وظيفته كعداد .

٤ -لقاء محاضرات تشرح التعليمات المختلفة التي تتعلق بعمل العدادين في الميدان . وتشمل هذه التعليقات كيف يمكن ان يكون العداد ناجحاً في

مقابلة المستجوب ، وشرح للاسئلة المختلفة التي تتضمنها الاستمارة ، الغرض من كل سؤال ومعناه وكيف يمكن توجيهه الى المستجوب وكيفية الحصول على الاجابات الصحيحة وكيفية تسجيل هذه الاجابات في الاماكن المخصصة لها، وكيفية مواجهة المشاكل المختلفة التي يمكن ان تعترض العداد اثناء عمله، وكيفية اجراء مراجعة سريعة لما كتبه في الاستمارات قبل مغادرة مكان المستجوب، وكيفية كتابة التقارير اليومية عن عمله، والاجراءات الادارية التي يتطلبها عمله في الميدان مثل استلام الاستمارات وعدها، وكيفية الاتصال بالمشرفين عليهم وكيفية تسليم الاستمارات بعد الانتهاء منها، وكيفية مواجهة النفقات التي يحتمل ان يحتاج الى انفاقها اثناء متابعة العمل .

ومناقشة هذه التعليمات تكون بشرحها واعطاء أمثلة كثيرة تكشف عن الصعوبات المختلفة التي يحتمل ان تظهر عند العمل بحيث تؤدي المناقشة الى فهم العدادين للتعليمات فيها دقيقاً وكاملاً واستيعابها بحيث تصبح شيئاً عادياً بالنسبة لهم وبحيث يكونوا مستعدين للاختبار فيها في أي وقت دون سابق انذار .

• - اختبار العدادين في نهاية التدريب ويتوقف استمرار العداد في عمله على نتيجة هذا الاختبار .

٦ - تدرب العدادين الذين اجتازوا الاختبار النظري في الميدان والقصد من هذا التدريب التحقق من أن العداد قد استوعب فعلاً التعليمات التي درسها نظرياً وأصبح بذلك قادراً على تطبيقها عملياً . كذلك يترتب على هذا التدريب ان يتعرف العداد على منطقة العد المخصصة له - حدود المنطقة والوحدات الواقعة فيها وأماكن وجود هذه الوحدات (أين تقع ابواب المساكن مثلاً) -، ويكون هذا التعرف على أساس مطابقة ما هو مسجل في القوائم الموزعة عليهم مع ما هو موجود على الطبيعة . ويمكن أن يجري هذا النوع من التدريب أثناء اجراء العد الاختباري الذي سيأتي مناقشته فيما بعد.

رابعاً - حصر وحدات المجتمع موضوع البحث :

في دراسة واسعة النطاق كتعداد السكان ، او التعداد الصناعي ، او التعداد الزراعي ... الخ . يكون من الضروري تقسيم الدولة الى مناطق عد يخصص كل منها لعدد واحد أو لفريق من العدادين . ويستلزم هذا العمل تحديد أقسام الدولة علي خرائط كبيرة للمناطق الريفية وعلى تصميمات هندسية للمدن . فاذا لم تكن هذه الخرائط والتصميمات متوفرة يكون من الضروري العمل على اعدادها خصيصاً لهذه المناسبة .

ويساعد العد الاختباري في تحديد الوقت الكافي للملا الاستمارات وهذا بالتالي يمكن أن يساعدنا في تحديد مساحة مناطق العد . وهذه المسألة من الأهمية بمكان كبير في تعداد السكان حيث يجب أن تجري عملية التعداد خلال وقت قصير جداً . فاذا كانت مناطق العد واسعة جداً فان عملية التعداد لا بد وان تنهار ، أما اذا كانت ضيقة جداً فان العدادين سوف يجدون أنفسهم بدون عمل خلال جزء من الوقت المحدد وبذلك نكون قد أسرفنا في الانفاق .

بعد تحديد مناطق العد تعد قوائم بأسماء وحدات الاستبيان في كل منطقة منها لتوزيعها على العدادين . وفي التعداد الصناعي تتخذ هذه القوائم شكل كشوفات بأسماء المؤسسات الصناعية وعناوينها ، وفي التعداد الزراعي تتخذ شكل كشوفات بأسماء الحائزين وعناوينهم ، وفي تعداد السكان تتخذ شكل كشوفات بأسماء وحدات التعداد التي قد تكون اسرا أو مؤسسات سكنية . ونلاحظ انه في هذا التعداد لا يكفي فقط اعداد هذه الكشوف بل يجب أيضاً ترقيم المساكن في المدن والقرى واعداد علامة مميزة يمكن وضعها عليها عند الانتهاء من العد بالنسبة لها .

ان العناية بعملية حصر وحدات الاستبيان واعداد كشوفات بها تحدد لنا ما إذا كان التعداد سوف يكون شاملاً أو جزئياً . وتظهر أهمية الحصر في تعداد السكان حيث لا يكون من المنطق ترك عملية تحديد موقع وحدات

العد للعدادين كي يقوموا بها اثناء قيامهم بعملية العد حيث يكون الوقت المحدد لهذه العملية ضيق جداً وبذلك لا يمكن أن يضيع العداد وقته بحثاً عن وحدات الاستبيان التي يطلب اليه زيارتها . ولذلك يجب أن يتم هذا العمل مقدماً حتى يستطيع كل عداد أن يتعرف على وحدات العد التي تدخل في منطقة العد الخاصة به قبل الوقت المحدد للتعداد .

أما في أنواع التعداد الأخرى قد يكون من الممكن أن نجتمع بين تحديد وحدات العد وزيارتها في نفس الوقت ، وبذلك يمكن تلافي النفقات التي يتحملها الجهاز الاحصائي في اعداد قوائم الحصر ، خاصة إذا تذكرنا أن العد في هذه التعدادات يكون على مراحل . الا انه من الأفضل اجراء الحصر قبل بدء التعداد حتى يمكن تدقيق عمل العدادين على الاقل ، بالاضافة إلى انه يمكن الاسترشاد بقوائم الحصر في تحديد عدد الاستمارات بدقة حتى تتجنب الاسراف . ومن ناحية أخرى قد يكون من المقرر اجراء التعداد لنوع معين من وحدات المجتمع الاحصائي موضوع البحث وبذلك تكون عملية الحصر ذات أهمية كبيرة حيث يمكن خلال هذه العملية معرفة الصفة الخاصة بوحدات المجتمع والتي نسترشدها في تحديد الوحدات التي يشملها التعداد وتلك التي لا يشملها ، مثلاً اذا تقرر أن يشمل التعداد الصناعي المؤسسات التي توظف ١٠ عمال وأكثر ، ففي هذه الحالة لا بد أن نعرف توزيع المؤسسات تبعاً لعدد المشتغلين فيها مقدماً حتى نقتصر فيما بعد أثناء عملية العد على المؤسسات المطلوبة إذ يطلب إلى العدادين زيارة هذه المؤسسات مباشرة تبعاً لعناوينها التي تظهر في كشوفات الحصر .

كذلك في الدراسات بالمعينة يكون من الواجب حصر وحدات المجتمع لاعداد قوائم باسمائها وعناوينها وبعض المعلومات المبدئية البسيطة عنها . وتساعدنا مثل هذه القوائم على تحديد اطار المعينة (ويجب أن يكون شاملاً دون أي ازدواج حتى لا تكون العينة متحيزة) ، وعلى تحديد

نوع العينة الذي يناسب المجتمع موضوع البحث إلا انه قد يتقرر سحب العينة على أساس وحدات كبيرة (العينة على مراحل) أو يتقرر أن يكون الاطار في شكل خريطة أو صور شمسية ، وفي هذه الحالات لا يكون هناك داعياً لحصر جميع وحدات المجتمع وانما نكتفي بحصر الوحدات الكبيرة ثم الوحدات الصغيرة في عينة المرحلة الاولى فقط (حصر الحيازات في عينة القرى مثلا) أو بأعداد الخرائط أو التصاميم أو الصور الشمسية التي يكون قد تقرر استعمالها كاطار للمعينة .

خامساً - تنظيم الدعاية :

مهما حاولنا أن نعد العدة للبحث الاحصائي الذي تجمع معلوماته من الجمهور فان نجاحه يتوقف في النهاية على استجابة المستجوبين بالتعاون في اعطاء المعلومات الصحيحة . ولهذا فالتنظيم المحكم لحملة من الدعاية لكسب ثقة الجمهور وتأييده يكون جزءاً هاماً من العمل التمهيدي .

ويرتكز جزء هام من كسب ثقة الجمهور وتعاونهم على العدادين أنفسهم ، ولهذا فتدريبهم في هذه الناحية له أهمية كبيرة . كذلك يكون لكسب تأييد المختارين وقادة الجمهور المحليين ورجال الدين أهمية كبيرة في كسب ثقة الجمهور . كل هذه الوسائل لها قيمتها ويجب استغلالها إلى أقصى حد ممكن . ولكن مع ذلك يجب الالتجاء إلى طرق الدعاية الشعبية لاهميتها في تعريف أفراد الجمهور باغراض الاستقصاء وأهدافه وفي أعدادهم لاستقبال هذا العمل بروح ودية .

والهدف الاول لحملة الدعاية يجب أن يكون هو تعريف الناس باغراض التعداد فتحاول بكل الطرق الممكنة ان نجعلهم يدركون الفوائد التي سوف تستخدم في تحقيقها المعلومات التي سوف يمطونها ، وبمعنى آخر كيف يكون لذلك أهمية كبيرة في وضع سياسة الحكومة نحو المسائل الاجتماعية

والاقتصادية وكيف تؤدي هذه السياسية إلى رفع مستوى المعيشة في الدولة إذا كان هذا هو الهدف .

وإذا ما استطعنا ان نفنسر أغراض التعداد أمكننا أن ننقل إلى الخطوة التالية وهي توضيح طريقة اجرائه . فيجب أولاً توضيح مسئولية الجمهور في اعطاء المعلومات المطلوبة (القانون) والوسيلة التي سوف تلجأ اليها الحكومة لجمعها ، ولا يجب أن يقتصر التوضيح على عملية العد فقط بزعم انها هي التي تتصل بالجمهور اتصالاً مرتبطاً وثيقاً ، ولكن يجب أن يشتمل التوضيح على الخطوات الاخرى التي سوف تمر بها المعلومات حتى يتم نشرها .

ولا يجب أن تتركز الدعاية حول الناحية السلبية فقط أي عدم استخدام المعلومات فيما يضر بصالح الأفراد ، بل يجب أن نحاول توضيح الجانب الايجابي كذلك يجعل أفراد الجمهور يدركون ان المعلومات تجمع لصالحهم . وفي معالجة هذه الناحية لا يجب أن نصوّر للناس آمالاً نعلم انها كاذبة حيث أن ذلك يسيء اليهم وإلينا في الأجل الطويل ، فقد ننجح في ذلك حالياً ولكن لا بد وأن يتولد لدى الجمهور الشعور بعدم الثقة ، الأمر الذي يجعل من الصعب أن نكسب تعاونه في المستقبل .

وللدعاية الشعبية نوعان رئيسيان من الوسائل هما الصحافة والاذاعة متضمنة التلفزيون كذلك . ويمكن الاعتماد اساساً على المقالات في الصحف اليومية لشرح جميع الوسائل المتعلقة بالدراسة . ولكن نظراً لقلة انتشار الصحف في الدول العربية يكون الراديو والتلفزيون هما اكثر الوسائل فعالية خاصة بالنسبة للذين يقطنون في المناطق الزراعية .

والنوع الشائع من الدعاية بواسطة الاذاعة هو سلسلة من الاحاديث عن الدراسة . ولا بد ان تكون هذه الاحاديث قصيرة وأن تحتوي على ما يثير اهتمام الجمهور . كذلك يمكن تنظيم برامج يُسأل فيها احد المسؤولين في الجهاز القائم بالدراسة أسئلة يوجهها أفراد يمثلون طوائف الجمهور المختلفة بحيث

تكون الإجابة عليها توضيحاً للنواحي المختلفة التي يقوم عليها التعداد . كذلك يمكن تنظيم تمثيلات بسيطة توضح فيها المسائل العملية المختلفة المتعلقة بالتعداد . ويجب ان نلاحظ ان تنظيم برامج الدعاية بالراديو او التلفزيون هو من اختصاص الفنيين ولهذا يجب ان يترك العمل في تعداد واسع النطاق الى قسم الدعاية الذي يقوم بالعمل فيه موظفون أخصائيون في هذه الناحية .

وفي حالة الدراسات التي تجري بواسطة المعاينة والتي تقتصر على مناطق محدودة ، كذلك في حالة الدراسات التي تكون طائفة قليلة من الجمهور هي المعنية بالدراسة (التعداد الصناعي مثلاً) يكون من الاسراف توجيه حملة الدعاية للسكان عامة ولهذا يمكن الاقتصاد على النشرات التي ترسل الى الطوائف المعنية فقط وعلى المقالات التي تنشر في الجرائد والمجلات المتخصصة . وفي التعداد الزراعي حيث تكون وحدات التعداد مبعثرة في جميع انحاء الدولة لا يمكن الاعتماد كثيراً على الصحف الا اذا وجدت الصحف المحلية وهو ما لا يتوفر في الدول العربية . وفي هذه الحالة تكون الاذاعة بالاضافة الى النشرات واللافتات التي تعرض في أسواق القرى أكثر وسائل الدعاية فعالية .

ومن الواضح ان توقيت الدعاية مسألة ذات أهمية كبيرة - فاذا ظهرت المقالات في الصحف واذيعت البرامج في الراديو والتلفزيون قبل التعداد بأسابيع طويلة فان اثرها لا بد ان يضيع عند حلول موعد التعداد وبذلك لا تحقق الفائدة منها . وبذلك يحسن ان نبدأ الدعاية في شكل بسيط للغاية ثم تقوى تدريجياً حتى بدأ التعداد . وفي تعداد السكان يجب ان تصل حملة الدعاية الى ذروتها خلال الايام القلائل التي تسبق يوم التعداد حتى يمكن بحلول هذا اليوم ان يكون كل فرد في الدولة قد اصبح لديه فكرة واضحة عما يجري حوله في هذا اليوم . أما في الدراسات التي يجري العد فيها على مراحل فان حملة الدعاية يمكن ان تستمر اثناء المدة الطويلة التي يجري فيها العد وخاصة في المناطق التي لم يشملها العد بعد وتلك التي يكون العد جارياً فيها .

وبالرغم من ان حملة الدعاية تهدف اساساً الى اثارة اهتمام الجمهور وكسب
تعاونه عند اجراء عملية العد ، الا انه لا يجب ان نوقفها عند الانتهاء من هذه
العملية ، فما دمننا قد اثرا اهتمام الجمهور فلا بد ان نحافظ على هذا الاهتمام
بنشر النتائج التي نتوصل اليها في صيغة شعبية بسيطة ، ويساعد ذلك في
ابقاظ الوعي الاحصائي لدى الجمهور ، الأمر الذي يساعدنا في اجراء الدراسات
التالية في المستقبل

ويجب ان نتذكر دائماً ان الدعاية عملية باهظة التكاليف ، ولهذا يجب
محاولة اقناع اصحاب الصحف على تقديم الفراغ المجاني لنشر المقالات القصيرة
واخبار التعداد . واذا لم ننجح في ذلك يكون من المهم عمل ميزانية دقيقة
للدعاية على اساس المفاضلة بين فعالية الوسائل المختلفة التي يمكن الالتجاء اليها
ولا يمكن ان ننجح في ذلك الا اذا تركنا امر الدعاية للفنيين فيها وسوف
نرى في النهاية ان ما ينفق على توظيفهم يبرره ما يستطيعون تحقيقه
من وفر في تكاليف الدعاية من جهة ومن زيادة فعاليتها من جهة اخرى .
واذا لم نترك امر الدعاية للفنيين فاننا لا بد ان نسرف في الاتفاق دون ان
نحقق النتائج المرجوة .

سادساً - القيام بعد اختباري :

بعد الانتهاء من جميع الخطوات التمهيدية لاجراء العد يجب اجراء عد
اختباري للتأكد من ان جميع العمليات التنظيمية التي اتخذت قادرة على تحقيق
وظائفها . والفائدة الأولى من العد الاختباري هي اختبار الاستمارة في الميدان قبل
القيام بالعد الفعلي إذ مهما دققنا في تصميمها فان الاطمئنان اليها لا يمكن أن
يتحقق إلا بتجربتها في الميدان للكشف عن عيوب في صياغة الألفاظ
أو في صياغة الأسئلة ، اذ قد يتبين لنا من الاختبار أن بعض الألفاظ المستخدمة
ليست من الواضح الذي تصوره في بادية الأمر ، أو ان بعض الأسئلة قد
فهمها المستجوبون فهمًا يختلف عن المقصود منها ، وبذلك يكون من الضروري

تغيير مثل هذه الألفاظ ومثل هذه الأسئلة اذا كان ذلك ممكناً أو تزويد
العدادين بتعليمات أكثر تفصيلاً لمساعدتهم في شرح الأسئلة للمستجوبين وفي
الحصول على الاجابات المطلوبة .

كذلك يساعدنا العد الاختباري في اختبار التنظيم الاداري الذي اعدناه
لاجراء الدراسة ، فقد يتبين لنا مثلاً أن الاجراءات التي اتخذت لتوزيع
الاستمارات فيها ضياع كبير للوقت وان تغيير بسيط في هذه الاجراءات
يمكن أن يوفر جزءاً من الوقت الضائع . كذلك قد يتبين لنا أن بعض
العدادين بالرغم من نجاحهم في الاختبارات النظرية لا يصلحون للقيام بالعمل
في الميدان بسبب الحجل والارتباك عند مواجهة الناس . كذلك يساعدنا
العد الاختباري في تحديد الوقت اللازم لكل وحدة استبيان وبذلك يكون
من الممكن تقدير ساعات العمل الضرورية لاجراء عمليات العد ، واذا اخذنا
في الاعتبار الوقت الذي يضيع في الانتقال من وحدة الى أخرى أمكننا أن
نتحقق من عدد العدادين الذين يجب علينا توظيفهم لاجراء عملية العد . ولا
شك أن كل هذه الاختبارات تساعدنا في التحقق من تقديراتنا للتكاليف
والتأكد من المبالغ التي خصصناها لاجراء الدراسة .

وعند اجراء العد الاختباري يحسن اختيار الوحدات التي يحتمل أن
تواجه معها أكبر قدر من المشاكل ، وفي نفس الوقت يحسن أن نجني بعض
الخبرة في معالجة المشاكل التي يمكن أن نواجهها في مختلف أنواع الحالات .
ولهذا لا يجب أن نعتبر العد الاختباري عملية صغيرة تجري في منطقة واحدة
نختارها لهذا الغرض ولكن كعملية واسعة النطاق تجري في أماكن تمثل
مناطق الدولة المختلفة .

ونطاق العد الاختباري يتوقف بلا شك على نوع الدراسة التي نريد القيام
بها ، فمثلاً في تعداد صناعي يمكن أن يكون العد الاختباري شاملاً لعدد من
المؤسسات في كل صناعة للتأكد من صلاحية الاستمارة لمختلف أنواع الصناعات

أما في دراسة ضيقة النطاق كبحث ميزانية الأسرة يمكن أن يتخذ العد الاختباري شكل اختبار دقيق للاستمارة بين بعض الأسر الشبيهة لتلك التي التي يتضمنها البحث وفي مثل هذه الدراسات ما يعيننا في الواقع هو اختبار الاستمارة وليس التنظيم الإداري الذي لا يكون له أهميته الكبيرة التي تظهر في حالة الدراسات الواسعة النطاق كتعداد السكان .

سابعاً - جمع المعلومات في الميدان

تعتبر عملية التعداد المرحلة النهائية التي تنتهي بها جميع الخطوات التمهيدية التي سبق القيام بها والتي ناقشنا أهمها فيما سبق. وفي تعداد السكان تجري عملية العد النهائية في وقت قصير جداً وهي لذلك تتضمن مجهوداً إدارياً ضخماً . وفي الأنواع الأخرى للدراسات الاستقصائية يمكن أن تجري عملية العد خلال وقت طويل ولهذا لا يكون لمشاكل التنظيم الإداري نفس التعقيد الذي نواجهه في تعداد السكان . وحتى في مثل هذه الحالات لا بد من الرقابة المحكمة على عملية العد والا استمرت طويلاً الأمر الذي يؤدي الى ضياع الوقت والمال .

وإذا ما تم تجنيد العدادين وتدريبهم ، وإذا ما تم تحديد مناطق التعداد وتنظيم قوائم بوحدات العد التي تتبع كل عداد ، تكون المشاكل التي نواجهها بعد ذلك أثناء اجراء عملية العد هي المواصلات والرقابة على العمل .

المواصلات :

يجب نقل كل تعداد الى منطقة اختصاصه بحيث يتواجد فيها عند بدء عمل كل يوم ، كذلك يجب اتخاذ الاجراءات اللازمة لتأمين انتقاله من كل وحدة الى الأخرى اذا استدعى الامر ثم انتقاله في نهاية اليوم الى منزله أو أي مكان يتفق عليه . وفي المدن الكبيرة حيث تكون وحدات التعداد متجاورة يكون الانتقال بينها على الاقدام وبذلك لا تنشأ مشكلة نقل العدادين بين

وحدات التعداد. أما في التعداد الزراعي أو أي استقصاء يتضمن عد وحدات متناثرة لا بد من اتخاذ الاجراءات اللازمة لتأمين انتقال العدادين بين الوحدات المختلفة . والوسيلة المناسبة للمواصلات تتوقف بالطبع على جغرافية المكان الذي يشمل التعداد وفي التعداد الزراعي الذي أجري في العراق استخدمت جميع وسائل المواصلات من القوارب للانتقال في المستنقعات الجنوبية والجمال والبغال التي استخدمت في المناطق الشمالية البعيدة . أما في تعداد السكان اذا استثنينا المناطق القروية البعيدة التي تكون فيها الاسر متناثرة فالغالب أن العدادين يستطيعون الانتقال بين الوحدات المختلفة بواسطة الاقدام وبذلك تقتصر مشكلة المواصلات في هذه الحالة على تأمين نقل العدادين الى أماكن اختصاصهم في الوقت المناسب .

وإذا كانت وسائل المواصلات العامة منظمة تنظيمًا جيدًا يكون كل ما يحتاجه هو تزويد العدادين بتصاريح تسمح لهم بالسفر المجاني مستخدمين هذه الوسائل (التابعة للبلدية أو الحكومة) . أما في الأماكن التي لا تصل إليها هذه الوسائل العامة والتي يجب على العدادين زيارتها يكون واجباً على الادارة الاحصائية أن تنظم وسائل المواصلات الخاصة بها وهذا اما بتأجير العربات الضرورية أو شرائها ثم بيعها عندما ينتهي العمل . وإذا كان التعداد يستغرق وقتاً طويلاً يكون شراء العربات هو أفضل الطرق التي يمكن أن تلجأ إليها الادارة الاحصائية لتأمين وسائل المواصلات من الناحية الاقتصادية حيث أن التأجير يكون في التالب أكثر تكلفة . وقبل تأجير أو شراء أي من وسائل المواصلات يجب بحث امكانية استخدام الوسائل العامة بحثاً دقيقاً شاملاً لان ذلك يكون حتماً أقل تكلفة

مراقبة عملية التعداد :

مهما كان مستوى العدادين عالياً ومهما كان تدريبهم دقيقاً فان الاشراف والرقابة على عملية العد من الأهمية بمكان كبير . ولهذا يجب التأكد من تنظيم

مناطق التعداد تنظيمياً يساعد على امكانية الاشراف والرقابة عليها أثناء اجراء عملية العد . ونوع الاشراف يختلف بطبيعة الحال تبعاً لنوع الاستقصاء الذي يطلب اجراؤه . وفي استقصاء واسع النطاق مثل تعداد السكان والتعداد الزراعي قد يكون أمراً ضرورياً تنظيم سلسلة مفصلة من الرقابة تبدأ من المكتب المركزي حتى تصل الى العداد في الميدان . وعدد الحلقات التي تشملها تلك السلسلة يتوقف على نطاق الاستقصاء وعلى درجة شموله لمناطق الدول المختلفة . وفي أي استقصاء واسع النطاق على كل حال ، لا يمكن تجاهل التنظيم المحلي .

وعندما تكون الادارة المحلية منظمة تنظيمياً جيداً وبالأخص في حالة الدراسات الاستقصائية الواسعة النطاق مثل تعداد السكان يحسن أن يوكل أمر الادارة الى السلطات المحلية تحت الاشراف الفني للادارة الاحصائية . وتعداد السكان يؤثر أحياناً على القانون والنظام في بعض المناطق ولذلك يحسن منح المسؤولية بطريقة يمكن بواسطتها تلافي أي ارتباك في الادارة المحلية . وحتى لو لم تجعل الادارة المحلية مسؤولة مباشرة عن اجراء التعداد يكون من المهم جداً اجراء التنظيم بالتعاون معها اذ يمكن أن يكون للاحتكاك معها آثار وخيمة بالنسبة لنجاح التعداد بينما يكون للتعاون الصادق معها أثره الحميد في كسب تعاون الجمهور وتجنب الصعوبات السياسية .

وعدد المشرفين الذين نكون في حاجة اليهم يتحدد تبعاً لعدد العدادين الذين يمكن للمشرف الواحد الاشراف عليهم . والمشرفون على العمل في الميدان يعلنون تقاريرهم الى المسؤولين المحليين ، فإذا كان عدد المناطق ليس كبيراً يمكن هؤلاء أن يبلغوا تقاريرهم إلى الادارة المركزية وبذلك يكون عملهم حلقة اتصال بين هذه الادارة وبين القائمين بالعمل في الميدان . وإذا كانت السلطة المحلية هي المسؤولة عن اجراء التعداد يكون من الواجب تحديد شخص واحد في كل ادارة يكون مسؤولاً عن الاتصال بالادارة المركزية للتعداد .

ويجب على كل مشرف أن يكون قادراً على الاتصال المتواصل بالعدادين أثناء قيامهم بالعمل في الميدان ، كما يجب تنظيم الاجراءات التي يستطيع بمقتضاها العدادون الاتصال بالمشرفين عليهم عندما يواجهون صعوبات تحتاج إلى العلاج السريع . وبذلك تكون أول وظيفة للمشرف هي أن يكون موجوداً في مكان معين ومستعداً للاجابة على الأسئلة المعقدة التي يمكن أن تظهر أثناء العمل ولمساعدة العدادين في مواجهة المستجوبين الذين يمتنعوا عن اعطاء المعلومات المطلوبة .

والوظيفة الثانية للمشرفين هي التأكد من قيام العدادين بعملهم على الوجه الأكمل . ففي بعض الحالات عرف ان العدادين يقومون بملأ الاستمارات بأنفسهم وهم جالسون في المقاهي ، والاشراف الدقيق يجب أن يعمل على منع هذا من الحدوث ، وان كان التدقيق في اختيار العدادين وتدريبهم التدريب الحسن هو أفضل طريقة لتلافي أي تراخي يمكن أن يحدث .

والوظيفة الثالثة للمشرفين هي مراقبة تقدم العمل حتى يمكن التأكد من اتمامه في الوقت المحدد . فقد يظهر ان الوقت المحدد لاجراء التعداد في منطقة ما كان أكثر من اللازم بينما الوقت الذي حدد لمنطقة اخرى أقل من اللازم وفي مثل هذه الحالة يصبح من الضروري تحويل بعض العدادين من المنطقة الأولى الى الثانية وان كان يجب أن نتجنب مثل هذا الوضع بكل ما يمكن من الوسائل .

ولمساعدة المشرفين في هذه الوظيفة يجب على كل عداد أن يسجل ما قام به من عمل أولاً بأول ، ويمكن اجراء ذلك بتزويد العداد بقائمة بالوحدات التي يتحتم عليه زيارتها ويطلب منه ان يؤثر على كل وحدة ينتهي من جمع المعلومات عنها مع بيان الملاحظات التي تمن له بخصوصها . وفي نهاية اليوم تجمع هذه التقارير وتسلم للمشرفين حتى يمكنهم مراقبة تقدم العمل ، وبذلك يمكن التعرف على العدادين الذين يتسم عملهم بالبطء والذين يواجهون حالات كبيرة من عدم الاجابة فيوجه اليهم رقابة أكثر حتى يمكن استبدالهم إذا لزم

الأمر . فهناك أشخاص لا يمكن أن يكونوا عدادين أكفاء حيث لا يستطيعون تجنب معاداة المستجوبين ، وهذه الحقيقة قد لا تبدو واضحة إلا عند ما ينزلون إلى الميدان حيث تظهر عدم كفاءتهم في كثرة عدد حالات عدم الاجابة التي يبلقونها إلى المشرفين عليهم . ومثل هؤلاء العدادون يجب تحويلهم إلى أعمال أخرى بأسرع ما يمكن .

وفي نهاية كل يوم يجب تسليم الاستمارات التي تم ملأها الى المشرفين الذين يقومون بمراجعتها مع التقارير المقدمة من العدادين . فاذا وجد أن عدد الاستمارات التي انتهى منها أقل بكثير من العدد المقدر في بادئ الأمر ، يكون ذلك دليلاً على عيب في التنظيم العام للاستقصاء أو في الاستمارة لم تنتبه اليه هيئة التعداد . وفي هذه الحالة يجب أن تخطر الادارة المركزية بالأمر حتى يمكن بحث أسباب هذا العيب وعلاجه إذا أمكن ، وفي الدراسات الاستقصائية التي تستغرق عملية العد فيها وقتاً طويلاً يكون من الضروري مراقبة اجراء المراحل الأولى من العمل رقابة محكمة حيث يكون هناك متسع من الوقت لعلاج أي عيب في التنظيم أو في الاستمارة بشرط التنبه إلى هذه العيوب في وقت مبكر . ولا يكون هذا ممكناً في الدراسات التي تتركز فيها عملية العد في وقت قصير وبذلك ينحصر جهد المشرف في التأكد من شمول العد لجميع الوحدات في الوقت المحدد إذ لا يكون هناك وقت لعلاج أي عيب يظهر في التخطيط الأصلي .

ويجب مراجعة الاستمارات المنتهية مراجعة سريعة في نهاية كل يوم للتأكد من عدم وجود إجابات غير منطقية على بعض الأسئلة ومن الاجابة على جميع الأسئلة . وهذه المراجعة المبدئية في الميدان ذات أهمية كبيرة حيث يكون من السهل اتخاذ الاجراءات اللازمة لزيارة المستجوبين مرة ثانية إذا وجد في اجاباتهم أخطاء واضحة ، بينما لا يكون ذلك سهلاً إذا وصلت الاستمارات إلى الادارة المركزية دون التنبه الى مثل هذه الأخطاء في الميدان . كما أن معلومات العاملين في الميدان المحلي عن المستجوبين لها فائدة كبيرة في المساعدة في

اجراء مثل هذه المراجعة والتي لذلك يجب أن تعتبر جزءاً هاماً من أعمال الميدان .

ثانياً - تدقيق المعلومات

ان عملية مراجعة البيانات التي جمعت في الاستمارات للتأكد من صحتها قبل اجراء التبويب عليها تسمى بعملية تدقيق الاستمارات ، وتهدف هذه العملية الى تحقيق الامور الآتية :

١ - التأكد من ان البيانات التي جمعت صحيحة ودقيقة الى اكبر درجة ممكنة .

٢ - التأكد من تماسك البيانات فيما بينها وعدم تعارضها مع الواقع .

٣ - التأكد من تسجيل البيانات في الاستمارات تبعاً للتعليمات الخاصة بذلك

٤ - التأكد من شمول الاستمارات لجميع البيانات المطلوبة .

٥ - التأكد من تسجيل البيانات تسجيلاً يساعد على التبويب .

٦ - التأكد من تنظيم البيانات في الاستمارة تنظيماً يساعد على ترميزها وتبويبها .

٧ - ملاحظة تعليقات العدادين والمشرفين عليهم حيث يمكن أن يكون لهذه التعليقات فائدتها في تفسير نتائج الدراسة .

ويجب أن نشير قبل مناقشة عملية تدقيق الاستمارات الى ان جمع المعلومات يجب أن ينظم بحيث لا يكون هناك حاجة إلى تصحيح أو تغير في البيانات التي جمعت الا في الحالات النادرة فقط ، فاقصداً الوقت يكون من الأفضل منع الخطأ بدلاً من وقوعه ثم محاولة تصحيحه . وقد سبق أن ناقشنا الوسائل المختلفة التي يمكن بواسطتها منع الخطأ قبل اجراء الدراسة الاختبارية وتجنيده

العدادين تبعاً للمستوى المطلوب وتدريبهم التدريب الكافي والاشراف عليهم
في الميدان .. الخ ..

خطوات تدقيق الاستمارات

في الدراسات الواسعة النطاق تبدأ عملية تدقيق الاستمارات في الميدان
حيث يجب أن يطالب العدادين بمراجعة الاستمارات التي ينتهون منها قبل
تسليمها للمشرفين عليهم ، ويحسن أن تتم هذه الخطوة في مكان المستجوب
وقبل مغادرته ، وعند استلام المشرفين الاستمارات يجب أن يتأكدوا من
الآتي :

١ - وجود اجابة لكل سؤال في الاستمارة ، حتى ولو كانت هذه
الاجابة مجرد تسجيل علامة معينة .

٢ - كتابة الاجابات بخط واضح يمكن قراءته دون خطأ .

٣ - ان الاجابات سجلت في أماكنها طبقاً للتعليمات الخاصة بها .

٤ - التماسك بين اجابات الأسئلة المختلفة ، بمعنى التأكد من عدم وجود
تعارض بين العمر والحالة الزوجية مثلاً ، بين المهنة والحالة التعليمية ، بين
عمر الام وعمر الابن البكر ، بين عدد العمال في المؤسسة ونتاجها ، بين
مجموع مساحات الحيازة المستخدمة في الاغراض المختلفة والمساحة العامة
للحيازة ، بين دخل الأسرة ومجموع انفاقها اذا اشتملت الاستمارة على اسئلة
تبين الادخار والاقتراض .. الخ .

٥ - إذا ظهر عدم تماسك بين اجابات بعض الأسئلة ، هل تنبه العداد
إلى ذلك وهل سجل في الملاحظات تفسيراً لهذا التعارض .

٦ - ان البيانات المسجلة تعكس بدقة اجابات المستجوبين .

٧ - عدم وجود أي خطأ في أسماء وعناوين المستجوبين كما هي مسجلة في الاستمارات .

٨ - وجود امضاء العداد وتاريخ الانتهاء من الاستمارات .

٩ - عدد الاستمارات في حوزة كل عداد وعدد الذي سلمه منها .

ويجب ان تجري عملية المراجعة من قبل المشرفين أولاً بأول ، أي يجب عليهم مراجعة كل ما تسلموه من الاستمارات دون انتظار انتهاء عملية جمع المعلومات ، إذ بذلك يمكن تنبيه العدادين الى أخطائهم حتى لا تتكرر معهم في الاستمارات التالية . كما ان هذه المراجعة يمكن أن تنبه المشرفين الى سوء فهم العدادين لبعض الأسئلة أو التعليقات الخاصة بها أو لكيفية توجيه الأسئلة أو لكيفية تسجيلها في الاستمارات وبذلك يكمل تصحيح ما فهموه خطأ أثناء تدريبهم كذلك فان المراجعة بهذه الطريقة يمكن أن تساعد العدادين على تذكر اجابات المستجوبين نظراً لعدم مضي وقت طويل على مقابلتهم وبذلك يمكن تصحيح بعض الأخطاء دون الحاجة إلى إعادة المقابلة مرة ثانية ، الأمر الذي يمكن أن يضايق المستجوبين كثيراً . كذلك فان المراجعة بهذه الطريقة يمكن ان تنبه المشرفين الى العدادين الذين يستمرون في تسليم استمارات فيها نفس الأخطاء التي وقعوا فيها أولاً وبذلك يمكن استبعادهم من أعمال الميدان واحلال آخرين محلهم .

كذلك فان المراجعة بهذه الطريقة يمكن أن تكشف عن وجود نقص في التعليقات فيما يختص بتفسير بعض الأسئلة أو بعض المصطلحات المستخدمة أو بعض التعاريف ، الأمر الذي يساعد على معالجة هذا النقص في وقت مبكر قبل أن تنتهي أعمال الميدان ويصبح من المستحيل اجراء العد مرة أخرى .

والخطوة التالية هي مراجعة الاستمارات في المكتب القائم بالدراسة مراجعة تفصيلية ، ذلك لأن مراجعة الميدان نظراً للسرعة التي تجري بها تكون

مراجعته عامة لا يمكن الاعتماد عليها كلية ، أو تقتصر فائدتها على اكتشاف السهو أو الأخطاء الواضحة البسيطة حتى يمكن تصحيحها في الميدان عندما تكون الاستثمارات لا زالت قريبة من المستجوبين . ويكون الهدف من المراجعة في المكتب التأكد من أن البيانات التي جمعت ليس بها أي تعارض يمكن أن يكشف عن أخطاء نتجت عن أي سبب من الأسباب .

وتنشأ مشكلة إدارية عند إجراء المراجعة المكتبية ، وهي هل يقوم بمراجعة الاستثمار الوحدة مراجع واحد ، أو هل من الأفضل تقسيم الاستثمار الى أقسام يقوم بمراجعة كل منها مراجع خاص ؟ ولا شك أن الطريقة الثانية تكون هي الأفضل إذا كانت استثمار البحث طويلة ومعقدة حيث أنها تساعد المراجع على التخصص في ناحية معينة يستطيع أن يجيدها بالبحث والمران . إلا أن التوسع في تقسيم عملية المراجعة يمكن أن يكون خطراً حيث أنه لن يساعد على الكشف عن عدم التماسك بين البيانات في الأقسام المختلفة ، الأمر الذي لا يمكن أن يكشف إلا بالمقارنة بين إجابات الأسئلة المختلفة في الاستثمار . على أنه يمكن الجمع بين الطريقتين في المراجعة وذلك بتكليف مراجع واحد بمراجعة الاستثمار بأكملها من حيث التماسك بينما يقوم مراجعون آخرون متخصصون بمراجعة بعض الأرقام في الاستثمار التي تحتاج الى مراجعة خاصة . ومن المستحسن إذا كانت الاستثمار تحتوي على بيانات تحتاج مراجعتها الى إجراء عمليات حساب أن تكون من اختصاص مراجع مزود بآلة حاسبة تحقيقاً للسرعة والدقة في العمل .

ولكي يستطيع المراجع تدقيق أي بيان لا بد أن يكون لديه علم كافٍ بالموضوع حتى يستطيع أن يقبل الإجابات غير المحتملة أو المفتعلة أي التي يكتبها العدادون في الاستثمارات دون مقابلة المستجوبين المعنيين . ويكون هذا العيب بالغ الخطورة في الدراسات المتعلقة بأراء المستجوبين حيث لا تكون البيانات مصورة لأرائهم بل لأراء العدادين، الأمر الذي يحكم على الدراسة بالفشل الذريع وتلافياً لهذا العيب يمكن مطالبة العدادين بتدوين تعليقات

المستجوبين وملاحظاتهم ، ولا شك أن العدادين لن يكونوا من الذكاء والمهارة بحيث يستطيعوا تصور ملاحظات وتعليقات تختلف فيما بينها اختلافاً جوهرياً وبحيث تبدو مقبولة من المراجعين . وتزداد هذه المشكلة تعقيداً اذا قابل العدادون مستجوبين غير الذين حددوا لهم حيث تأتي تعليقاتهم وملاحظاتهم متباعدة ومقبولة ، الأمر الذي لا يمكن أن يكشفه المراجعون . والوسيلة الوحيدة لاكتشاف مثل هذه الاجابات هي التأكد من عمل العدادين في الميدان ثم اجراء اختبار على الاستمارات بارسال مراجعين لمقابلة عينة من المستجوبين مرة ثانية .

وفي بعض الأحيان يظهر عدم دقة الاجابات من مراجعة البيان الخاص بمجموعة من المستجوبين وليس بواحد منهم فقط ، ويكون ذلك بمقارنة المجموع الخاص ببيان ما بما نتوقعه ، فاذا لاحظنا فرقاً كبيراً كان ذلك دليلاً على أن المستجوبين لم يتحروا الدقة عند الاجابة عن السؤال الخاص بهذا البيان . ومثلاً على ذلك ما حدث في تعداد السكان عام ١٩٠٠ في أمريكا حيث تبين بعد التبيوب أن عدد الزوج الذين لا يستطيعون الكلام باللغة الانجليزية كبير إلى درجة توحى بالشك . وقد ظهر من تدقيق الاستمارات أن ترتيب الأعمدة الخاصة بالأسئلة - القدرة على القراءة باللغة الانجليزية والقدرة على الكتابة بها والقدرة على التحدث - أدى إلى هذا الخطأ ، حيث كانت الأعمدة متلاصقة فكان العداد نتيجة للسرعة يكتب لا ، لا ، لا . في الوقت الذي كان يجب أن يكتب فيه لا ، لا ، نعم . ويتضح من هذا المثال أن هذا الخطأ لم يكن من الممكن أن يكشف بمراجعة استمارة كل مستجوب على حدة ، أما بمراجعة الاستمارات ككل اتضح الخطأ ، الأمر الذي ساعد على تصميم الاستمارة تصميماً مختلفاً عند اجراء التعداد التالي .

تماسك الاجابات

ان من أهم عناصر عملية تدقيق الاستمارات التأكد من عدم وجود تعارض

بين اجابات الاسئلة المختلفة ويكون ذلك باجراء المقارنة بين اجابات الأسئلة المرتبطة ببعضها طبقاً لنظام منصوص عليه في التعليمات الخاصة بالمراجعين . وفي بعض الأحيان يكون من الممكن تحديد الخطأ في اجابة أي من الأسئلة التي ظهر التعارض فيما بين اجاباتها وبذلك يمكن اجراء التصحيح في المكتب دون الرجوع إلى المستجوب . وفي حالات أخرى لا يمكن التحديد وبذلك يجب الرجوع إلى المستجوب لاجراء التصحيح وإذا كان مستحيلاً يجب تصنيف اجابة هذا المستجوب ضمن فئة « لم يبين » .

وغالباً تكون عملية مقارنة الأرقام التي تتضمنها الاجابات المختلفة عملية منفصلة يقوم بها مراجع خاص تكون وظيفته التأكد من ان الجاميع تتطابق مع الأجزاء التي تكونها ، وان كل رقم يقع ضمن نطاق المنطقة الخاص بالبيان المتعلق به ، وان كل رقم لا يتعارض مع أي رقم آخر في الاستمارة ، وان الأرقام التي تظهر في استمارات العدادين لا تختلف اختلافاً كبيراً عن الأرقام الموجودة في استمارات العدادين الآخرين (في الدراسات الخاصة بطبقة معينة من المجتمع) . فاذا تبين وجود تعارض في أي حالة من هذه الحالات لا بد من التحقق من الأرقام أو من وجود تفسير بذلك عند العدادين سجلوه في ملاحظاتهم على ما جمعه من بيانات .

وحدة القياس :

بالرغم من أن الاجابات قد تكون سليمة الا انه لا يمكن تبويبها إلا بعد تحويلها إلى وحدات قياس موحدة . فإذا كان العدادون قد لاحظوا التعليمات الخاصة بذلك عند جمع البيانات فان البيانات سوف تكون مسجلة على أساس موحد ، ولكن مع ذلك يجب اجراء المراجعة للتأكد من ان الامر هو هكذا . كذلك فان بعض المستجوبين يمكن أن يكونوا قد اجابوا بتفصيل غير مطلوب فتكون وظيفة المراجعين شطب التفاصيل حتى لا يختلط الامر على المرزبن مثلاً اجاب المستجوب على السؤال الخاص بالعمر فاعطى سنه

بالسنوات والشهور والايام ، او أجاب عن السؤال الخاص بدخله فأعطى رقماً باللبرات والقروش وسجل العداد الاجابة كما لفظها المستجوب . كذلك يجب التأكد من أن طريقة تسجيل الاجابات موحدة بين الاستمارات المختلفة ، مثلاً استعمال العلامة - لتدل على عدم انطباق السؤال على المستجوب بالنسبة لبعض العدادين واستعمالها لتدل على انطباق السؤال لعدادين آخرين .

وعموماً يمكننا أن نلاحظ أن وحدة القياس التي جمعت وسجلت تبعاً لها البيانات في الاستمارات المختلفة يمكن أن تتحقق إذا كانت التعليمات الخاصة بها واضحة وإذا تمسك العدادون بهذه التعليمات عند العمل في الميدان ولاحظها المشرفون عند اشرافهم على العدادين .

السهو :

يمكن أن تأتي الاستمارات الى المكتب وبها بعض السهو أي بها اسئلة لا يظهر أمامها ما يدل على ان العداد قد وجهها الى المستجوبين وحصل على اجابة خاصة بها ، ولعل ذلك هو السبب - كما أشرت سابقاً - إلى أن التعليمات يجب أن تنص على تسجيل اجابة عن كل سؤال في الاستمارة ويمكن أن تكون الاجابة مجرد تسجيل علامة معينة . كذلك يمكن أن يكون السهو فيما يختص بتسجيل جزء من الاجابة ، مثلاً عدم وضع علامة الليرة بجانب الرقم الخاص بالدخل . وعلامة الفدان بجانب الرقم الخاص بالمساحة المزروعة بأي محصول من المحاصيل ... الخ ... مثل هذه الاستمارات سوف تهمل عند اجراء الترميز ، الأمر الذي يضخم فئة لم تبين الى درجة كبيرة والذي يجعل المطلعون على نتائج الدراسة لا يعولون كثيراً عليها . لذلك يكون واجب المراجعين التأكد من عدم وجود أي سهو في الاستمارات ، وكثير من الاجابات الناقصة يمكن معرفتها في المكتب باستنتاجها من اجابات الأسئلة الاخرى مثلاً السؤال هل تملك سيارة ؟ نعم - لا ، ثم السؤال ما هي ماركة السيارة ؟ فاذا وجدت اجابة خاصة بماركة السيارة أمكن معرفة أن اجابة

السؤال الأول هي نعم. على أن بعض الاجابات غير المسجلة يكون من الصعب تحديدها في المكتب ولما كان من المستحيل في بعض الاحيان الرجوع الى المستجوبين لذلك يجب مطالبة المشرفين في الميدان بالتنبه الى هذا العيب في الاستمارات عند جمعها من العدادين ، وبذلك عندما ترد الاستمارات الى المكتب لا يجب أن يظهر فيها هذا العيب الا نادراً .

كذلك يجب أن يتأكد المراجعون من عدد الاستمارات ، والواقع أن هذه العملية يحسن أن تكون من اختصاص موظف خاص يقوم بعد الاستمارات عند استلامها من الميدان ، إلا ان ذلك لا يمنع المراجعين من اعادة عد الاستمارات بعد الانتهاء من مراجعتها وقبل تسليمها لقسم الترميز حيث يمكن أن يختفي بعض منها في قسم المراجعة نتيجة سهو احد المراجعين أو بعض منهم. كذلك يجب على المراجعين التأكد من استبعاد أي نسخة من الاستمارات حيث ان استمرار وجود هذه النسخ عند الترميز والتبويب فيما بعد يؤدي الى أخطاء جسيمة في النتائج ، وتوجد هذه النسخ أحياناً عند مطالبة العدادين باعادة كتابة استمارة ما أو بعض الاستمارات بسبب كثرة الاخطاء أو سبب عدم وضوح الخط المكتوب به البيانات وتسليم النسخ مع اصولها لقسم المراجعة مرة ثانية .

قبول أو رفض الاستمارة :

من وظائف المراجعين أن يقرروا ما إذا كانت استمارة ما دقيقة وكاملة بحيث يمكن قبولها للتبويب او انها ليست كذلك فيجب رفضها . وتنشأ هذه المشكلة بشكل واضح عند الدراسة بالمعينة حيث يجب التأكد من ان الاستمارات تحتوي على البيانات الخاصة بالوحدات التي سحبت في العينة وليس بوحدات اخرى . ولا شك ان اكتشاف هذا الخطأ يكون من وظيفة قسم مراقبة عملية المعينة ، إلا ان المراجعين يجب أن يتأكدوا من عدم وجود هذا الخطأ قبل اجراء التبويب .

ويحسن أن تبين التعليقات إلى المراجعين المستوى الذي يمكن على أساسه قبول أو رفض الاستمارة وعلى المراجعين أن يتقيدوا بهذا المستوى ولا يجيدوا عنه . على أن رفض بعض الاستمارات يؤثر كثيراً على نتائج الدراسة ولهذا يحسن أن يعاد النظر في أمر هذه الاستمارات من قبل رئيس قسم المراجعين والخير في المعاينة (إذا كانت الدراسة بالمعاينة) .

اعادة تنظيم بنود الاستمارة :

لا يكون ترتيب الاسئلة في الاستمارة ، احيانا ، متفقاً مع الترتيب الضروري لتسهيل عملية الترميز والتبويب ، لهذا تنقل البيانات إلى استمارات اخرى بالترتيب المطلوب لقسم التبويب . على أن هذه الخطوة يحسن تلافيتها نظراً لاضاعتها للوقت ولاحتمال الخطأ عند النقل . ويمكن الاستعاضة عن ذلك باعادة ترقيم الاسئلة بلون واضح وبارز .

التعليات الى المراجعين :

يجب أن تطبع التعليقات الى المراجعين في كتيب خاص يكون جاهزاً قبل اجراء العمل في الميدان . ويحسن أن يرفق بهذا الكتيب نماذج من استمارات فيها امثلة من الاخطاء والعيوب التي يحتمل أن تظهر لتكون مرشداً عملياً لاجراء عملية المراجعة . ويمكن أن يحتوي كتيب التعليقات على الآتي :

١ - على المراجعين ان يكونوا ملمين إماماً كاملاً بالتعليات الموجهة الى العدادين والمرمزين بجانب المامهم بالتعليات الخاصة بهم حيث ان ذلك يساعدهم على اكتشاف الاخطاء المختلفة .

٢ - لا يجب على المراجعين بأي حال من الأحوال ان يقضوا بالمسح او الشطب على معالم البيان الاصلي المسجل في الاستمارة عند تصحيحه .

٣ - يجب استخدام قلم من لون آخر عند إجراء التصحيح .

٤ - لا يجب إجراء أي تصحيح إلا إذا وجد ما يبرر ذلك من إجابات الأسئلة الأخرى في الاستشارة .

٥ - يجب تحديد البيانات التي يمكن للمراجعين تغييرها والبيانات التي لا يجب عليهم تصحيحها إلا بعد أخذ رأي رئيس القسم .

٦ - يجب استشارة رئيس القسم فيما يختص بالاستشارات التي يرى المراجعون استبعادها . ولا يجب وضع هذه الاستشارات في الملف الخاص بها إلا بعد أخذ رأي المشرف على عملية المعاينة (في حالة الدراسات بالمعاينة) .

٧ - يجب وضع العلامة الخاصة بالمراجع عند تصحيح أي بيان حتى يتمكن من مراجع بعده أن يعرف سبب هذا التصحيح .

٨ - يجب وضع التعليقات الخاصة بكيفية إجراء التصحيح أو التغيير ، مثلاً شطب البيان بخط أفقي لإلغاء هذا البيان إلغاء تاماً والشطب بخطوط رأسية إذا أريد إلغاء جزء من البيان وشطب البيان وإجراء التصحيح فوقه أو تحته في مكان واضح إذا أريد التصحيح .

٩ - يجب تنبيه المراجعين إلى أن علمهم يمكن أن يراجع مرة ثانية ولذا يجب أن يكونوا مستعدين لتفسير أي تغيير أو تصحيح أجروه على الاستشارات .

١٠ - يجب إثبات التاريخ على الاستشارات المنتهية من المراجعة وكذلك توقيع المراجع حتى يمكن مراقبة عمل المراجعين من ناحية السرعة والدقة .

كذلك يجب أن يشتمل كتيب التعليقات على تفسير خاص بكل بيان في الاستشارة ، كيف يكون البيان مقبولاً وكيف لا يكون كذلك إلا إذا كان مصحوباً بتفسير خاص . كذلك يجب أن توضح التعليقات أجزاء الاستشارة التي يراجع تماسكها مع الأجزاء الأخرى . وحيث أن البيانات تكون مرتبة في الاستشارة لتسهيل عملية جمع المعلومات في الميدان ، لذا يكون من المرغوب فيه

أن توضح التعليقات افضل ترتيب للمعلومات عند إجراء المراجعة إذا كان هذا الترتيب يختلف عن الترتيب الموجود في الاستشارة . كذلك يحسن ان تُظهر التعليقات البيانات التي يحتمل ان تظهر فيها الاخطاء وكيفية اكتشافها عند المراجعة . ومن الأهمية بكان كبير أن تبين التعليقات الأسس التي بناء عليها يمكن رفض الاستشارات أو قبولها لإجراء عمليات التبويب عليها ، وتعلق هذه الأسس غالباً بشمول الاستشارة للبيانات المطلوبة وصحة هذه البيانات ودقتها ، وكذلك بصحة اعتبار الاستشارة ممثلة لوحدة في العينة اذا كانت الدراسة تجري بالمعينة .

كذلك يجب أن يساعد تنظيم عملية المراجعة على حصول كل مراجع على نسخة من أي تعديل أو تغيير في التعليقات الخاصة بهذه العملية . إلا انه يجب أن نلاحظ ان التعليقات الخاصة بعملية المراجعة لا بد من إعدادها بحيث لا يجري عليها أي تغيير إلا في الحالات القصوى فقط حيث ان تعديل التعليقات من حين لآخر يؤدي الى ضياع الوقت ، إذ يترتب على هذا التعديل إعادة مراجعة الاستشارات التي انتهت مراجعتها ، كما ان المراجعين يفقدون ثقتهم واحترامهم للشرفين عليهم ، فضلاً عن فقدان حماسهم للعمل وتقاعسهم عن الاستفسار إذا واجهتهم أية مشكلة ، وبذلك يفقد المشرفون على المراجعة سلطتهم وقدرتهم على ضبط العمل .

ومن الأهمية بكان كبير ان يفهم المراجعون ان لهم الحق في الاستفسار إذا واجهتهم أية مشكلة فنية ، ولا بد أن يكون هناك تنظيم خاص بذلك حيث يمكن ان تكتب هذه الاستشارات على بطاقات مقاس ٣ × ٥ بوصة وعند ورودها إلى رئيس قسم المراجعة تصنف في ملفات خاصة حسب نوع البيان أو حسب أقسام الاستشارة مع الاجابات والتوضيحات الخاصة بها، وبذلك يمكن الرجوع اليها قبل توضيح أي استفسار جديد حتى لا يحدث أي تعارض فيما بينها . وإذا قدم الاستفسار من عدة مراجعين يحسن أن يصدر رئيس

القسم نشره تفسيرية كي توزع على الجميع . ولا بد من الاحتفاظ بهذه الملفات كي تسلم بعد إنهاء عملية المراجعين إلى القائمين بكتابة التقرير على الدراسة حتى تكون مرشداً لهم عند تحليل البيانات والتعليق عليها .

تنظيم قسم المراجعة :

في الدراسات الضيقة النطاق يمكن ان يقوم مراجعان بالعمل ، ومن الخطأ أن يعتقد المشرف على الدراسة انه يستطيع أن يشرف على جمع المعلومات وفي نفس الوقت يقوم بالمراجعة حيث انه عند جمع المعلومات تنشأ مشاكل كثيرة وملحة لا تترك وقتاً كافياً للمشرف كي يقوم بعملية المراجعة في الوقت المناسب .

وفي الدراسات الواسعة النطاق يجب ان يكون هناك قسم خاص للقيام بعملية المراجعة ويجب أن يعمل هذا القسم على أساس يمكن به ضبط العمل من ناحية السرعة والدقة ، ولتحقيق ذلك يجب أن يحدد متوسط عدد الاستمارات التي يمكن مراجعتها في وحدة زمن معينة ، كما يجب أن يميز عمل كل مراجع حتى يمكن تحديد المراجعين المتراخين في العمل والذين لا يتمشى عملهم مع مستوى الدقة المطلوب .

وغالبا ينشأ الاحتكاك بين المراجعين والعدادين ، الأمر الذي يوقع المشرفين على عملية المراجعة في مشاكل كثيرة . وينشأ هذا الاحتكاك لأسباب كثيرة منها ان العدادين يعتقدون ان مطالبة المراجعين لهم بتوخي الدقة التامة في تسجيل البيانات ليس إلا أمراً نظرياً ومظهر من مظاهر محاولة المراجعين إظهار سلطتهم . كذلك يعتقد العدادون أن عملهم في الميدان أكثر صعوبة من عمل المراجعين الذين يجلسون في مكاتبهم ثم يرفضون استمارة ما لمجرد ان بيان ما لم يسجل فيها . وعلى رئيس قسم المراجعة ان يمنع هذا الاحتكاك من التوسع حتى لا يكون ذلك سبباً من أسباب تعطيل العمل . ولهذا يجب أن

يدرب المراجعين على كتابة ملاحظاتهم بشكل موضوعي دون المساس بشخص
العداد نفسه ، وإحدى الوسائل لتحقيق موضوعية المراجعة أن يشير المراجع
عند تصحيح أي بيان إلى الصفحة والفقرة من التعليمات التي تتعلق بموضوع
التصحيح . كذلك يمكن تدريب العدادين على عملية المراجعة ، حيث ان هذا
التدريب ولو انه ليس الا تدريباً مقتضباً يجعل العدادين يقدرّون الحاجة إلى
الدقة في العمل ويعطيهم فكرة أفضل عما يمكن أن تكون عليه الاستمارة
مقبولة لإجراء عملية التبويب عليها فيما بعد . ويعتقد البعض ان هذا التدريب
للعدادين يجعلهم قادرين على اصطناع البيانات التي تنال رضى المراجعين ،
وهذا أمر خطر للغاية حيث ان الدراسة بذلك لا تصور واقع الأمر وإنما
تصور ما يتخيله العدادون من بيانات . على انه في الدراسات الواسعة النطاق
والتي تشمل الاستمارة فيها على أسئلة كثيرة لا يكون لدى العدادين الوقت
لاصطناع البيانات خاصة إذا كانت عملية الإشراف على العدادين والمراجعة
منظمة ودقيقة .

ويجب أن يكون عدد المراجعين في المكتب كافياً لمراجعة الاستمارات
بمجرد وصولها الى القسم القائم بالعمل حتى يمكن تنبيه العدادين الى أخطائهم
والاستفادة منها عند تكملة العمل في الميدان . وفي بعض الحالات تطالب
المكاتب المحلية بارسال بعض من الاستمارات المنتهية إلى قسم المراجعة لأخذ
فكرة عن مستوى العمل في الميدان ، فاذا أتبع هذا الاجراء يجب ان ينبه على
هذه المكاتب بعدم ارسال الاستمارات التي يعتقدون أنها ممتازة وأنه يجب أن
تلتفت عينة عشوائية من الاستمارات بجانب الاستمارات التي يعتقدون أن بها
صعوبات تحتاج لدراسة المتخصصين في عملية المراجعة .

وفي بعض الدراسات تنظم عملية المراجعة بحيث تجري محلياً في مناطق
الدولة المختلفة وحتى يمكن توحيد الأسس التي تجري تبعاً لها هذه العملية يسمح
لرئيس كل مكتب محلي أن يصدر التعليمات والتفسيرات للموظفين الذين يعملون
معه على ألا يعمل بها إلا بعد موافقة الرئيس العام للمراجعة في المكتب
المركزي .

الفصل الثالث

التصنيف والتبويب

تبين لنا من الفصل السابق أن المعلومات تجمع في استمارات كثيرة يصل عددها في بعض الدراسات الى عشرات الآلاف . والمعلومات بهذا الشكل لا يمكن أن توضح لنا اتجاهات الظواهر التي نريد دراستها ، بل لا يمكن بأي حال من الأحوال أن تساعدنا في اجراء التحليل الرياضي عليها والذي نرغب في اجرائه للتوصل الى المقاييس المختلفة التي نهذف اليها من وراء البحث .

نفترض اننا نريد دراسة وضع السكان بالنسبة للنشاط الاقتصادي ، فان احد اهدافنا من هذه الدراسة يكون معرفة عدد السكان الذين يكوّنوا القوة العاملة في الدولة متضمنة المشتغلين والمتعطلين وعدد السكان الذين لا يعملون ولا يرغبون في العمل لاسباب مختلفة (أما لأنهم طلبة أو ربات منازل أو أصحاب أملاك يكتفون بإيراداتهم من أملاكهم) وأخيراً عدد السكان الذين لا يستطيعون العمل (اما لأنهم أطفال صغار أو لأنهم مسنين فوق سن معين أو لانهم عجزة). اننا نوجه سؤالاً لكل فرد من السكان ونحصل منه على اجابة؛ الا ان هذه الاجابات المتفرقة في استمارات لا تساعدنا في الوصول الى هدفنا ، ولذلك يجب تصنيف هذه الاجابات في الفئات المختلفة التي نريدها ثم وضع نتائج التصنيف في جدول يظهر لنا الارقام المطلوبة ، هذه العملية نسميها بالتصنيف والتبويب .

وهناك أنواع مختلفة للتبويب ولكنها جميعاً تهدف الى هدف واحد هو تجميع المعلومات في فئات و ابرازها باكبر وضوح ممكن في أضيق حيز ممكن. وعهد اجراء هذه العملية لا يجب أن يكون غرضنا قاصراً على عرض البيانات عرضاً يتفق والبحث الذي نقوم به فقط ، خاصة اذا كان لهذه البيانات أوجه متعددة بحيث يمكن الاستفادة منها في أبحاث أخرى غير البحث الذي من أجله جمعت المعلومات (يحدث هذا غالباً في الابحاث الاحصائية الوصفية) . ولهذا يجب عرض البيانات عرضاً يساعد كل من يكون له غرض في البحث أن يجد ما يريد من معلومات تفيد . فالمعلومات الخاصة بالسكان في تعداد للسكان مثلاً ، يجب تبويبها بحيث يمكن أن يستفيد منها الباحث الاقتصادي والاجتماعي والبيولوجي ... الخ . فالاقتصادي يستطيع أن يلجأ اليها لمعرفة عدد المشتغلين في أوجه النشاط الاقتصادي المختلفة ، والاجتماعي يجد فيها ما يساعده على دراسة التركيب العمري والجنسي للسكان بجانب توزيعهم تبعاً للجنس المختلفة ، والبيولوجي يستطيع أن يعتمد عليها في دراسة نمو السكان والتغيرات التي يمكن أن تصيب التركيب العمري من فترة الى اخرى .

ان الغرض من التصنيف هو كما ذكرنا تجميع الوحدات في فئات معينة وتختلف انواع هذه الفئات تبعاً لطبيعة البيانات المطلوب تبويبها وتبعاً للكيفية التي على أساسها سوف تستخدم بعد تبويبها . وأهم انواع التصنيف هي التصنيف الزمني وفيه تصنف المعلومات تبعاً لوحداث زمنية معينة كالشهر أو السنة مثلاً ، والتصنيف الجغرافي وفيه تصنف المعلومات تبعاً لمناطق جغرافية معينة قد تكون هي الاقسام الادارية للدولة وقد تكون مناطق مختلفة في العالم ، والتصنيف الوصفي وفيه تصنف المعلومات تبعاً لوصاف معينة تشترك فيها كل مجموعة من الوحدات مثلاً تصنيف السكان تبعاً للجنس وتصنيف المؤسسات الصناعية تبعاً لنوع نشاطها ، والتصنيف الكمي وفيه تصنف المعلومات تبعاً لفئات رقمية معينة مثلاً تصنيف الطلبة تبعاً لفئات أعمارهم حيث يمكن وضع الطلبة الذين في سن العشرين في مجموعة وهؤلاء الذين

في سن الواحدة والعشرين في مجموعة اخرى وهكذا . ويمكن توسيع مدى المجموعة فنقول عدد الطلبة الذين تتراوح أعمارهم بين ١٦ - ٢٠ سنة هم ١٠٠ طالب وعدد الذين تتراوح أعمارهم بين ٢٢ - ٢٤ سنة هم ٥٥ طالب وهكذا . ويسمى هذا التوزيع بالتوزيع التكراري وسوف نناقشه بتفصيل فيما بعد .

هذه هي الانواع المختلفة لتصنيف المعلومات من الناحية النظرية ، إلا أنه من الناحية العملية يمكن تصنيف الوحدات تبعاً لأساسين من هذه الاسس ، فنصنف السكان مثلاً تبعاً لآعمارهم في مناطق الدولة المختلفة او نصنف المؤسسات الصناعية تبعاً لنوع نشاطها وفئات حجمها أو نصنف الطلبة تبعاً لفئات أطوالهم وفئات أوزانهم فنقول مثلاً ان عدد الطلبة الذين تتراوح أوزانهم بين ٦٠-٦٥ كيلو والذين في نفس الوقت تتراوح أطوالهم بين ١٥٠-١٥٥ سم هم ٣٠ طالب ، أو تصنف الاطفال تبعاً لفئات أعمارهم وفئات أوزانهم فنقول ان الاطفال الذين تتراوح أعمارهم بين ٢ - ٤ سنوات والذين تتراوح أوزانهم بين ١٢-١٥ كيلو هم ٥٠ طفلاً مثلاً . ومثل هذا النوع من التبويب يسمى بالتبويب المزدوج حيث يكون هناك أساسان لتصنيف الوحدات أي أن كل رقم في الجدول يدل على عدد من الوحدات ذات صفتين في نفس الوقف .

وليس هناك ما يمنع من تصنيف الوحدات تبعاً لأكثر من أساسين فنصنف السكان مثلاً تبعاً لمكان سكنهم في الريف أو المدن وتبعاً لجنسهم ، ذكور أو اناث ، وتبعاً لحالتهم الزوجية أعزب أو متزوج أو مطلق أو أرمل ، أو نصنف المواليد تبعاً لجنسهم وتبعاً لفئات أعمار امهاتهم وتبعاً لشرعيتهم - مواليد شرعيين أو غير شرعيين . إلا أنه يجب أن نلاحظ انه كلما زادت عدد الاسس التي تصنف تبعاً لها الوحدات يصبح التبويب الناتج من عملية التصنيف معقداً يصعب فهمه ، ولذلك يحسن دائماً ألا نكثر في أسس التصنيف إلا إذا كانت الحاجة الماسة تدعو الى ذلك .

طرق التصنيف .

يمكن اجراء التصنيف بطرق مختلفة ، واختيار طريقة منها يتوقف على عوامل كثيرة منها عدد الوحدات التي نريد تصنيفها ومدى تنوع المعلومات الخاصة بهذه الوحدات. فكلما كان نطاق الدراسة ضيقاً كلما احتجنا الى اجراء التصنيف باليد حيث لا نحقق فائدة تذكر من استخدام الآلات وحيث يكلفنا استخدامها تكاليف كثيرة اذا قيست بالنسبة لنطاق البحث . أما اذا اتسع نطاق البحث بحيث أصبح يشمل وحدات كثيرة ومعلومات مختلفة يكون استخدام الآلات أمراً واجباً حيث نحقق فوائد كثير من السرعة في انجاز العمل وانقاص اليد العاملة التي نكون في حاجة اليها . وتظهر هذه الفوائد بشكل واضح في تعداد السكان حيث نحتاج الى معرفة النتائج في وقت يمكن فيه الاستفادة منها وحيث يصل عدد الوحدات الى الملايين .

التصنيف اليدوي :

لتوضيح طريقة التصنيف اليدوي نفترض أن لدينا استمارات خاصة بخمسين مؤسسة صناعية ونريد تصنيفها تبعاً لموقعها . نبدأ بأعداد الجدول الذي تجري عملية التصنيف فيه ونسميه جدول التفريغ وطبعاً يوضع في كعب الجدول مناطق الدولة المختلفة - مثلاً بيروت ، لبنان الشمال ، لبنان الجنوب ، الجبل ، البقاع . وبعد ذلك نتناول الاستمارات الواحدة تلو الاخرى ونقرأ البيان الخاص بموقع المؤسسة ، فاذا كان بيروت نضع أمامها اشارة معينة لتكون / . هذه الاشارة تدل على وجود وحدة (في هذه الحالة مؤسسة) في هذه الفئة للموقع. ونكرر هذه العملية حتى ننتهي من تفريغ البيان الخاص بموقع الخمسين مؤسسة . ولتسهيل عملية عد هذه الاشارات يحسن أن نضع كل أربع اشارات يحوار بعضها هكذا ////، أما الاشارة الخامسة فتكتب بعكس اتجاه الاشارات الأربعة على هذا النحو 4444. وبذلك يتكون لدينا مجموعات من الاشارات كل

منها يحتوي على خمسة مثلاً // متباعدة عن بعضها نوعاً ما وبذلك يمكن عدّها بسهولة (١٧ مؤسسة) . بعد ذلك يصمم جدول تنقل فيه المعلومات الموجودة في جدول التفريغ ولكن بدلاً من الاشارات تكتب الارقام الدالة عليها وبذلك نكون قد أعددتا جدولاً يظهر توزيع المؤسسات تبعاً لموقعها .

ويمكن أن يجري العمل بنفس الطريقة في حالة التبويب المزدوج والفرق الوحيد ان الاشارة توضع أمام صفتين تدلان على الوحدة التي نصفها . ففي مثالنا السابق اذا كنا نصف المؤسسات تبعاً لموقعها وفئات حجمها نبدأ بأعداد جدول التفريغ ونكتب في كعبه فئات الموقع وفي عناوين أعدته فئات الحجم التي نتفق عليها مثلاً ١ - ٤ ، ٥ - ٩ ، ١٠ - ١٩ ، ٢٠ - ٤٩ ، ٥٠ - ٩٩ ، ١٠٠ أو أكثر . بعد ذلك نتناول الاستمارات الواحدة تلو الأخرى ونقرأ البيان الخاص بالموقع مثلاً بيروت ثم نقرأ البيان الخاص بعدد المشتغلين في المؤسسة مثلاً ٦٧ عامل فنضع الاشارة أمام بيروت وتحت فئة الحجم ٥٠ - ٩٩ وهكذا نستمر في وضع الاشارات بنفس الطريقة السابقة .

الا اننا قد نجد ان عدد الوحدات في كل استمارة من الاستمارات الخمسين كبيراً بحيث نضيع وقتاً طويلاً اذا أردنا تمثيل هذه الوحدات بالاشارات السابقة وفي هذه الحالة يحسن كتابة العدد نفسه الخاص بعدد الوحدات ونضيف اليه العدد في الاستمارة التالية وهكذا . فاذا أردنا تصنيف المؤسسات والمشتغلين فيها تبعاً لفئات الموقع فاننا نصمم جدول التفريغ حيث نضع في كعبه فئات الموقع ثم نقسم جسم الجدول الى قسمين ، قسم خاص بتفريغ المؤسسات وقسم لتفريغ المشتغلين . ثم نتناول الاستمارات الواحدة تلو الأخرى ونقرأ البيان الخاص بموقع المؤسسة فنضع الاشارة / أمام هذا الموقع ثم نقرأ البيان الخاص بعدد المشتغلين فتكتب العدد أمام نفس الموقع وهكذا . ويظهر ذلك في جدو التفريغ الآتي : -

توزيع المؤسسات والمشتغلين فيها تبعاً لموقع المؤسسة

عدد المشتغلين	عدد المؤسسات	الموقع
١٠+٨+٥+١٨+٣٧+٢٣+٢	// 444	بيروت
٢٨ + ٣٥ + ١٥	///	الشمال
٥ + ٢٢ + ٦٧ + ١٢	////	الجنوب
١٥ + ٩	//	الجبل
٣٠+٧٢+٤٣+٢٥+١٩+٨	/ 444	البقاع

ويتضح من جدول التفريغ السابق أن عدد الاشارات الخاصة بعدد المؤسسات يكون هو نفسه عدد الاعداد الخاصة بالمشتغلين وهذا أمر منطقي حيث أن كل اشارة تدل على مؤسسة وكل مؤسسة لا بد أن يكون فيها عدداً من المشتغلين . وواضح من الجدول اننا وجدنا ٧ مؤسسات في بيروت يشتغل فيها ٢ + ٢٣ + ٣٧ + ١٨ + ٨ + ٧ + ١٠ أي ١٠٥ مشتغل . وبذلك يكون متوسط عدد المشتغلين للمؤسسة الواحدة = $\frac{105}{7}$ مشتغل وهكذا بالنسبة لمناطق الدولة الأخرى .

ويمكن اتباع نفس الطريقة إذا أردنا أن نصنف المؤسسات والمشتغلين تبعاً لموقع المؤسسة ونوع نشاطها أو تبعاً لموقع المؤسسة وفئة حجمها ، وفي هذه الحالة يتضمن كعب الجدول الأساسان اللذان نستخدمهما في التصنيف كما يظهر من الجدول التالي :

توزيع المؤسسات والمشتغلين فيها تبعاً لموقع
المؤسسة وفئة حجمها (مقاساً بعدد المشتغلين)

عدد المشتغلين	عدد المؤسسات	الموقع وفئات الحجم
		بيروت
٢	/	١ - ٤
٥ + ٨	//	٥ - ٩
١٠ + ١٨	//	١٠ - ١٩
٢٣ + ٣٧	//	٢٠ - ٥٩
		٥٠ - ٩٩
		١٠٠ وأكثر
		الشمال
		—
		—
		—

التصنيف الآلي :

لاحظنا ان التصنيف اليدوي يؤدي الى ضياع الوقت والجهد ولذلك لا
يمكن الاعتماد عليه الا في البحوث التي يقل فيها عدد الوحدات وكمية البيانات

الخاصة بها . اما في البحوث الواسعة النطاق يتحتم علينا اجراء التصنيف باستخدام الآلات الاحصائية . واول ما يجب ان نلاحظه على الآلات الاحصائية انها لا تفهم البيانات الكلامية وانما تفهم الارقام فقط وفي شكل ثقب . واستعمال هذه الآلات يقوم على الاسس الآتية :

١ - الاستعاضة عن الاستارة ببطاقة خاصة من الورق المقوي يمكن ان تتداولها الآلات وتجري عملها عليها .

٢ - التعبير عن الوحدة بثقب معين امام رقم معين وفي عمود معين من البطاقة ، فاذا كنا نريد تصنيف السكان تبعاً للجنس - ذكور واثاث - وخصصنا العمود الثالث في البطاقة مثلاً لهذا البيان وعبرنا عن فئة الذكور برقم ١ وفئة الاثاث برقم ٢ فاي ذكر من السكان سوف يعبر عنه بثقب في العمود الثالث امام رقم ١ في البطاقة . وواضح ان هذا الثقب يجري آلياً اي باستخدام آلة خاصة بهذه العملية .

٣ - فرز البطاقات بعد تثقيبها في آلة خاصة بذلك حيث تقوم هذه الآلة بتجميع البطاقات المثقوب فيها الرقم ١ في العمود الثالث مع بعضها وتجميع البطاقات المثقوب فيها الرقم ٢ في العمود الثالث مع بعضها :

٤ - توضع البطاقات بعد فرزها في آلة اخرى تقوم بعد البطاقات في كل مجموعة وتسجل لنا العدد امام رقم ١ اي الذكور وامام الرقم ٢ اي الاثاث .

٥ - واخيراً تؤخذ هذه النتائج المسجلة وتوضع في شكل جدول .

وبذلك عندما نقرر استخدام الآلات الاحصائية يتعين علينا اجراء عمليات معينة هي اعداد الدليل ، والترميز ، ومراجعة الترميز ، والتثقيب ، ومراجعة التثقيب ، والفرز والتبويب ، كما يتعين علينا استخدام الآلات التي تساعدنا في هذه الاعمال المختلفة (ابتداء من عملية التثقيب) .

اعداد الدليل :

يجري الترميز للبيانات بناء على دليل موضوع ومعد لهذا الغرض . والمقصود بالدليل القائمة التي تحتوي على الفئات المختلفة الخاصة ببيان ما والأرقام التي ترمز إلى هذه الفئات فنقول مثلا الدليل الخاص بالمهن أي القائمة التي تحتوي على المهن المختلفة التي يمكن أن توجد في مجتمع ما والأرقام الرمزية التي تدل على هذه المهن عند اجراء العمليات الخاصة بالتبويب . كذلك نقول الدليل الخاص بالنشاط الاقتصادي أي القائمة التي تشتمل على مختلف أنواع النشاطات الاقتصادية والأرقام الرمزية التي تدل عليها . كذلك نقول الدليل الخاص بالتقسيم الاداري للدولة أي القائمة التي تشتمل على التقسيمات الادارية المختلفة في الدولة والأرقام الرمزية الدالة عليها ... الخ .

ولاعداد دليل خاص ببيان ما يجب أن تتبع الخطوات التالية :

١ - يجب أن يكون واضحاً لدينا المعنى المقصود من البيان والغرض الذي نتوخاه احصائياً من السؤال عن هذا البيان . فإذا أردنا مثلاً أن نعد دليلاً للتقسيم الاداري للدولة يجب أن نفهم جيداً المقصود بهذا التقسيم وهل نحتاج في دراستنا إلى تقسيم اداري مفصل أو تقسيم عام قائم على أساس الأقسام الرئيسية في الدولة ، ويتضح من هذا المثال العلاقة بين الترميز والتبويب حيث أن البت في هذه المشكلة يتوقف على الجداول الاحصائية المعدة لتصنيف البيانات فيها . كذلك يتضح العلاقة بين الترميز وتصميم الاستمارة والتعليمات إلى العدادين حيث إذا كان المطلوب تصنيف البيانات تبعاً للأقسام العامة يمكن ان يكفي بتسجيل عنوان وحدة الاستبيان على أساس القسم الاداري الرئيسي ، أما اذا كان المطلوب التصنيف تبعاً للأقسام الفرعية يتحتم تسجيل العنوان مفصلاً . وإذا أردنا مثلاً أن نعد دليلاً للمهن يجب أن نفهم أولاً المقصود بالمهنة وهل نحتاج في دراستنا إلى تصنيف عام للمهن أو تصنيف للمهن على أساس عام ومفصل في نفس الوقت . وبالطبع يترتب على هذا البحث أن تؤخذ العلاقة بين التبويب والترميز والاستمارة والتعليمات العدادين في الاعتبار .

٢ - بعد أن نخلص من الخطوة الأولى نكون الخطوة الثانية هي حصر جميع الحالات التي يمكن أن نحصل عليها بتوجيه السؤال المعد في الاستمارة إلى جميع المستجوبين موضوع الدراسة . ويمكن أن يكون هذا أمراً سهلاً كما هو الحال بالنسبة للأسئلة الخاصة بالبيانات البسيطة مثل الجنس والحالة الزوجية والحالة التعليمية و .. الخ ، ويمكن أن يكون هذا الحصر أمراً صعباً كما هو الحال بالنسبة للمهن والنشاطات الاقتصادية والأمراض التي يمكن أن يصاب بها أفراد المجتمع موضوع الدراسة ونظراً لصعوبة اعداد الدليل الخاص بهذه البيانات قامت المنظمات الدولية المعنية بشؤون الاحصاء بوضع دليل لكل من هذه البيانات يمكن أن يعتمد عليه .

٣ - والخطوة الثانية هي وضع أرقام تدل على الحالات التي تم حصرها في الخطوة السابقة

مثال :

في دراسة عن الصناعة اشتملت الاستمارة على سؤال عن الكيان القانوني للمؤسسات الصناعية والمطلوب اعداد دليل خاص بهذا البيان . ان الخطوة الأولى كما قدمنا هي فهم المقصود بالكيان القانوني ثم حصر جميع الحالات الخاصة بهذا الكيان وهي :

مؤسسات فردية
مؤسسات تضامن
مؤسسات توصية بسيطة
مؤسسات توصية بالاسهم
شركات مساهمة الخ .

بعد ذلك يعطي كل من هذه الكيانات القانونية رقم خاص فيتكون لدينا قائمة بالدليل المطلوب يكون كالآتي :

- ١ مؤسسات فردية
- ٢ مؤسسات تضامن
- ٣ مؤسسات توصية بسيطة
- ٤ مؤسسات توصية بالأسهم
- ٥ شركات مساهمة

يعطي لنا هذا المثال فكرة بسيطة عن اعداد الدليل ، إلا اننا في هذه العملية يمكن أن نواجه مشاكل مختلفة تجعل العمل أكثر صعوبة . والمشاكل التي نواجهها في هذا الصدد يمكن تلخيصها في الآتي :

أولاً : صعوبة حصر جميع الحالات الخاصة ببيان ما ، وقد سبق أن أشرنا إلى هذه المشكلة فيما يختص بالمهن والنشاط الاقتصادي والأمراض وأسباب الوفاة والبضائع التي تدخل في التجارة الخارجية . لذلك أعدت اللجنة الاحصائية للمجلس الاقتصادي والاجتماعي بالامم المتحدة دليلاً للنشاط الاقتصادي وبضائع التجارة الخارجية - كما أعد مكتب العمل الدولي دليلاً للمهن ، وكذلك أعدت منظمة الصحة العالمية دليلاً للأمراض والابذاءات واسباب الوفاة . وكل دليل دولي يشتمل على حصر شامل للحالات التي يمكن أن توجد فيما يتعلق بالبيانات الخاصة بهذا الدليل ، وكذلك على ترميز لهذه الحالات . كذلك أعد كل دليل منها بحيث يمكن أن يستخدم على أساس أن المطلوب اجراء تصنيف تفصيلي يشتمل على الأقسام الرئيسية وما يتفرع عن هذه الأقسام من حالات فرعية وبذلك لا تضيق العلاقة بين هذين النوعين من التصنيف . كما ان كل دليل منها قد أعد بحيث يكون مرناً أي يمكن أن يتسع لادخال حالات جديدة إذا ظهرت مثل هذه الحالات في المستقبل وبذلك لا يستدعي الأمر تعديل الدليل من وقت إلى آخر .

وبالنسبة لهذه المشكلة يمكن أن أورد الملاحظات التالية :

أ - في بعض الأحيان لا يكون تفصيل الحالات الخاصة ببيان ما أمر

مرغوب فيه إذا اخذ في الاعتبار التبويب النهائي وبذلك يمكن اعطاء الحالات المطلوب ظهورها في التبويب ارقاماً رمزية ثم تجمع بقية الحالات في فئة واحدة دون تفصيل وتعطي جميعاً رقماً رمزياً واحداً . ففي مثالنا السابق إذا كان الاهتمام موجه إلى معرفة عدد المؤسسات الفردية والتضامن والمساهمة فقط فيكون الدليل كالآتي :

١ مؤسسات فردية

٢ مؤسسات تضامن

٣ مؤسسات مساهمة

٤ مؤسسات أخرى

وفئة المؤسسات الأخرى يقصد بها كل أنواع الكيان القايوني التي لم يخصص لها بند مستقل في الدليل . ونظراً لأنه من الصعب في بعض الأحيان تحديد جميع الحالات التي يمكن أن توجد فإن فئة أخرى مثل مؤسسات أخرى أو غراض أخرى تكون ذات فائدة قصوى حيث انها فئة مطلقة يمكن أن تستوعب كل الحالات التي لم يخصص لها بند مستقل أو التي لم يتوقع وجودها مصمم الدليل ، وبذلك يكون الدليل المنشأ مغطياً لجميع الحالات وهو أساس يجب أن ننظر اليه كحجر الزاوية عند انشاء اي دليل .

ب - وحيث ان كثيراً ما ترد الاستمارات إلى المكتب دون أن يستوفي بعض بياناتها لأي سبب من الأسباب ، ونظراً لصعوبة الرجوع إلى الميدان لاستيفائها - واطهار هذه الحالات في التبويب النهائي أمر بالغ الأهمية لأنه يمكن أن يكون مقياساً لمدى وعي المستجوبين ولمدى الدقة في جمع المعلومات - لذلك يجب أن يخصص في الدليل رقماً رمزياً للدلالة على كل بيان لم ترد الاجابة عنه ، وعادة يخصص آخر رقم في حدود الخانة لذلك الدليل مثلاً :

ذكر ١

انشى ٢

لم يبين ٩

وقد خصص الرقم ٩ لفئة لم يبين لأنه آخر رقم في خانة الأحاد وهي الخانة التي يتكون منها هذا الدليل ، وإذا كانت أرقام الدليل تتكون من خانتين فيخصص لفئة لم يبين الرقم ٩٩ لأنه آخر رقم في حدود خانة العشرات .

ج - يمكن أن يكون التفصيل في الدليل أمر مرغوب فيه ولكن يجب أن يؤخذ في الاعتبار عند انشاء دليل مفصل يحتوي على عدد كبير من الحالات عدد الخانات التي سوف يشغلها الدليل بناء على هذا التفصيل نظراً لأن لعدد الخانات التي يشغلها الدليل أهمية كبيرة من الناحية العملية التطبيقية وكذلك من ناحية بطاقة الآلات التي سوف ينقل اليها الدليل بالثقيب . وقد تكون هناك رغبة في تقليل عدد الحالات الخاصة بدليل ما تجنباً للتعقيد في العمل عند مطالبة المرمزين بتطبيقه وتوفيراً للوقت والجهد والنفقات بتقليل عدد الخانات في بطاقة الثقيب وما يلي ذلك من خطوات حتى اخراج التبويبات بواسطة الآلات ... الخ . لهذا يجب أن تقارن دائماً بين الفائدة التي تتحقق من التفصيل في الدليل وبين الجهد والوقت والنفقات وتعقيد العمل الذي يتطلبه تطبيق دليل مفصل . ولتجنب التفصيل في دليل ما ، تعطى الحالات الاساسية أرقاماً رمزية وتجمع الحالات الاخرى في فئة واحدة تسمى بحالات اخرى مثل حالة (مؤسسات اخرى) التي سبق الإشارة اليها .

د - كذلك لا بد عند وضع الدليل أن نراعي توفر المرونة فيه - وقد سبق أن أشرت الى ذلك فيما يختص بالتصنيفات الدولية . ويقصد بالمرونة أن يكون في الدليل موضع لاضافات جديدة دون أن تحدث تلك الاضافات قلقلة في مبنى الدليل . على انه يجب أن نلاحظ أن البيانات التي يوضع لها

دليل تختلف فيما بينها من ناحية حاجتها إلى المرونة فهناك بيانات بطبيعتها لا تستدعي ذلك حيث لا مجال فيها لأي حدث ، ومثال ذلك البيان الخاص بجنس السكان (ذكر أو انثى) . وهناك بيانات بطبيعتها عرضة لهذا المستحدث مثل البيان الخاص بالمهنة أو النشاط الاقتصادي أو التوزيع الجغرافي ... الخ .

مثال :

إذا أردنا وضع دليل للتقسيمات الادارية في مصر نلاحظ أن هناك محافظات ومديريات ويجب أن يوضح الدليل إذا كانت هذه المديريات في الوجه البحري أو في الوجه القبلي . فعند وضع الدليل لهذه التقسيمات الادارية يجب أن يتجه اهتمامنا الى المستقبل حيث يمكن أن تنشأ محافظة جديدة أو مديرية جديدة سواء في الوجه القبلي أو البحري ولذلك يجب أن يسمح الدليل الموضوع بادراجها فيه دون أن يهدم مبناه . وتطبيقاً لذلك يمكن أن يخصص العشرة أرقام الأولى للمحافظات وان كان عددها ٤ فقط ، ويخصص العشرين رقماً التالية لمديريات الوجه البحري وان كان عددها ٨ وبذلك يكون الدليل كالآتي :

اول محافظة	رقم ١	وآخر محافظة	رقم ٤
اول مديرية في الوجه البحري	رقم ١١	وآخر مديرية	رقم ١٨
اول مديرية في الوجه القبلي	رقم ٣١	وآخر مديرية	رقم ٣٨
اول أقسام الحدود	رقم ٥١	وآخر قسم	رقم ٥٥

فإذا أنشأت محافظة جديدة تعطى رقم ٥ وإذا أنشأت مديرية جديدة في الوجه البحري تعطى رقم ١١ وهكذا . وبالطبع نلاحظ أن المدى (عشرة أو عشرين أو أكثر) يتوقف في مثل هذه الحالة على عدد المفردات التي يتوقع زيادتها .

مثال ٢ :

في التصنيف الخاص بالمهنة أو بالنشاط الاقتصادي أو بالتجارة الخارجية يجري التقسيم على أساس أقسام وكل قسم يقسم إلى ١٠ مجموعات رئيسية مع أن بعض الأقسام يمكن أن لا تحتوي إلا على مجموعة واحدة أو مجموعتين فقط في الوقت الحاضر أما باقي الأرقام العشرة فتكون كفراغ يمكن أن توضع فيه أي مجموعة جديدة في المستقبل . وكذلك كل مجموعة رئيسية يمكن أن تقسم إلى عشرة مجموعات فرعية مع أنها في الحاضر قد لا تحتوي إلا على عدد أقل من عشرة مجموعات فرعية ويكون باقي الأرقام كفراغ توضع فيه أية مجموعات فرعية تنشأ في المستقبل .

وفي هذا الصدد نناقش مشكلة انشاء دليل لأقسام وفروع متفرعة عنها وتابعة لها . في هذه الحالة يخصص الرقم الأول للدلالة على القسم والرقم الثاني للدلالة على الفرع وإذا كان هناك ما يتفرع عن هذه الفروع يخصص الرقم الثالث للدلالة عليها . وبالنسبة لجميع الفروع المتفرعة عن قسم ما يكون دليلها جميعاً مشتركاً في الرقم الأول الدال على هذا القسم - مثلاً القسم رقم صفروما يتفرع عنه يكون الدليل الخاص به ٠٠ ، ٠١ ، ٠٢ ، ٠٣ . وهكذا ، والقسم رقم ١ وما يتفرع عنه يكون الدليل الخاص به ١٠ ، ١١ ، ١٢ ، ١٣ . وهكذا ، والقسم رقم ٢ وما يتفرع عنه يكون الدليل الخاص به ٢٠ ، ٢١ ، ٢٢ ، ٢٣ ؛ فإذا كان لدينا رقم دليل ٢٣ فيقرأ كآتي : الفرع رقم ٣ التابع للقسم رقم ٢ ، وإذا كان لدينا رقم دليل ٤٧ فنقرأه الفرع رقم ٧ التابع للقسم رقم ٤ ، وإذا كان لدينا رقم دليل ٥٨٢ فنقرأه الفرع رقم ٢ التابع للفرع رقم ٨ التابع للقسم رقم ٥ (يمكن الاطلاع على التصنيفات الدولية لأخذ فكرة تفصيلية عن هذا الموضوع) .

ولا شك أن ترك الفراغ في الدليل حق يمكن اضافة بنود جديدة إذا استحدثت مثل هذه البنود أمر يحقق التناسق في الدليل . ففي المثال الخاص

بالتقسيمات الادارية كان من الممكن أن تعطى المحافظات الارقام من ١ الى ٤ ومديريات الوجه البحري تبدأ بالرقم ٥ وتنتهي بالرقم ١٢ ومديريات الوجه القبلي تبدأ بالرقم ١٣ وتنتهي بالرقم ٢٠ وأقسام الحدود يبدأ الرقم ٢١ وتنتهي بالرقم ٢٥. وإذا استحدثت تقسيمات ادارية جديدة تعطي الرقم التالي ٢٦ سواء كانت تلك الاضافة محافظة أو مديرية او قسم حدود ، ولكننا بذلك نجعل الدليل غير متناسق وغير منطقي ، الامر الذي يجعل فهمه وتطبيقه أمراً صعباً . على أن ذلك لا يحدث إلا بالنسبة للبيانات المقسمة إلى أقسام تختلف فيما بينها ، إلا انه هناك بيانات ليست في حاجة الى مثل هذه المعالجة لأنه لا يترتب أي ضرر أن يدرج البند المستحدث في نهاية القائمة المعدة قبل استحداث هذا البند مثل البيان الخاص باسباب التعطل .

ثانياً : لا تنتهي عملية اعداد الدليل بحصر الحالات واعطائها أرقام تدل عليها ، إذ يجب توزيع البيانات على بطاقة التثقيب . والقاعدة الأساسية في هذا الصدد هي تخصيص عدد من الأعمدة في البطاقة لبيان ما تبعا لعدد الخانات التي يتكون منها الدليل الخاص بهذا البيان . وبذلك يتحدد مجموع الأعمدة التي يحتاجها تثقيب الأرقام الرمزية الخاصة بعدة بيانات جمعت في دراسة احصائية ما (تعداد السكان مثلا) وبالتالي يتحدد عدد بطاقات التثقيب التي نحتاجها ويتوقف ذلك طبعاً على نوع البطاقة المستخدمة ، ويتحدد هذا النوع تبعاً لنوع الآلات الاحصائية الموظفة في العمل فإذا كان مجموع الأعمدة التي نحتاجها ٨٠ أو أقل وكان عدد الأعمدة في البطاقة ٨٠ فاننا نحتاج بطاقة واحدة لكل وحدة ، أما إذا كان عدد الأعمدة التي نحتاجها أكثر من ٨٠ فاننا بذلك نحتاج لبطاعتين لكل وحدة وهكذا .

والمشكلة الثانية في هذا الصدد هي توزيع البيانات على بطاقة التثقيب ، بيان مثلاً يحتاج لعمودين هل يكون هذين العمودين في أول البطاقة أو في وسطها أو في آخرها . ويحتاج البت في هذه المشكلة الى خبرة فنية بالآلات الاحصائية المستخدمة ولا مجال هنا لبحث هذا الموضوع .

الترميز :

يقصد بعملية الترميز كتابة الرموز الرقمية أمام اجابات الأسئلة المختلفة طبقاً للدليل الذي أعدناه ، وبذلك اذا كانت بعض المعلومات في الاستمارة عبارة عن أرقام (عدد العمال مثلا) فلا يكون هناك حاجة لترميزها وانما يحسن احاطة الرقم بدائرة بلون خاص حتى يتنبه المثقوب اليها .

وهناك طرق مختلفة للترميز ، حيث يمكن أن تصمم اسئلة الاستمارة بحيث يكتب أمام كل سؤال الاجابات المحتملة له وكل اجابة لها رقم خاص بها يكون كرمز لها ، وبذلك عندما يؤثر المستجوب او العداد أمام احدى هذه الاجابات يكون في نفس الوقت قد حدد الدليل الرمزي الخاص بالاجابة ، وبهذه الطريقة تختصر اجراء عملية الترميز . بالرغم من الميزة الكبيرة لهذه الطريقة لا تتبع الا نادواً حيث أن اتباعها يحتاج الى تحديد جميع الاجابات المحتملة لكل سؤال تحديداً دقيقاً حتى لا تواجه اجابة ما ليس لها مكان في الاستمارة ، الأمر الذي يضطرنا في هذه الحالة الى اضافة فئة بعنوان « آخرون » الا أن هذه الفئة قد تتضخم بوقوع عدد كبير من الاجابات فيها وبذلك لا يساعد تبويب هذا البيان على اجراء التحليل الاحصائي الدقيق . وهناك عيب آخر لهذه الطريقة حيث ان اتباعها يستدعى استخدام فئات عامة وبذلك تضيع تفاصيل اجابة المستجوب ، الأمر الذي لا يساعد في اكتشاف أي خطأ وقع فيه العداد . واذا كان العداد مطالباً باجراء بعض العمليات الحسابية قبل تحديد اجابة المستجوب ضمن الفئات المطبوعة في الاستمارة فان أي خطأ يقع فيه العداد لا يمكن اكتشافه فيما بعد . هذا فضلاً عن أن خطأ العداد في وضع اجابة المستجوب أمام الفئة الخاصة به لا يمكن اكتشافه عندما ترد الاستمارات الى المكتب .

والطريقة الثانية للترميز هي ترميز البيانات على الاستمارة نفسها بعد جمع المعلومات ومراجعتها . واستخدام هذه الطريقة يساعدنا في توفير الوقت اللازم

لنقل البيانات بعد ترميزها الى كشف خاص بالترميز كما انه يمكن تقادي الأخطاء التي تحدث عند النقل . والعيب الأول لهذه الطريقة هو عدم توفر الفراغ اللازم لكتابة أرقام الدليل حيث أن الاستارة تسجل فيها المعلومات وملاحظات العدادين ثم يجري عليها التصحيح في الميدان ثم في المكتب، الأمر الذي يجعلها تكتظ بالكتابة . على أن هذه الصعوبة يمكن التغلب عليها بتخصيص فراغ خاص للترميز في الاستارة نفسها وتنبية العدادين الى عدم كتابة أي شيء في هذا الفراغ وكذلك يمكن أن يستخدم حبر بارز اللون عند اجراء الترميز . والعيب الاساسي لهذه الطريقة هو أن ترتيب البيانات في الاستارة قد لا يتفق مع ترتيب البيانات في بطاقات التثقيب حيث يكون للآلات أثر كبير في توزيع البيانات على أعمدة هذه البطاقات ، هذا فضلاً عن كثرة الاستارات المرمزة التي تذهب الى المثقبين وتكون سبباً في ارتباكهم في العمل .

والطريقة الثالثة هي ترميز البيانات في كشوفات تصمم خصيصاً لهذا الغرض ، ويمكن تصميم هذه الكشوفات بحيث ننقل البيانات اليها في شكل رموز رقمية تكتب في اعمدة معنونة بعناوين البيانات المختلفة ويمكن ترتيبت هذه الاعمدة بحيث تتفق مع الترتيب اللازم لاجراء التثقيب بسرعة وهو الترتيب الذي يتفق تماماً مع ترتيب الاعمدة في بطاقة التثقيب حسب تصميمنا لها .

والطريقة الرابعة هي ترميز البيانات في بطاقة ترميز خاصة بكل حالة على حدة ، وبذلك تختلف عن الطريقة السابقة حيث انه في هذه الطريقة يكون لكل وحدة بطاقة ترميز خاصة بها بينما في الطريقة السابقة ترمز البيانات الخاصة بعدة وحدات في كشف واحد ويتوقف عدد الوحدات على اتساع الكشف الذي نستعمله . والميزة الاساسية لهذه البطاقة انه يمكن استخدامها كبطاقة أرشيف نرجع اليها كلما اردنا تصنيف خاص للبيانات لم يجري من قبل .

والعيب الاساسي لهذه الطريقة هو ما نتحمله من تكاليف عندما نقرر استخدامها حيث نحتاج لعدد كبير منها خاصة في الدراسات الواسعة النطاق.

على أنه يمكن الجمع بين بعض هذه الطرق حيث تصمم الاسئلة ذات الاجابات المحددة بحيث تشتمل على الترميز الخاص بها على الاستمارة نفسها بينما تترك الاسئلة الاخرى دون ترميز على الاستمارة على أن يجري ترميز اجاباتها في المكتب بعد ذلك اما على الاستمارة نفسها أو في كشف ترميز مفصل أو في بطاقات ترميز .

ملاحظات على إعداد الدليل والترميز :

١ - يحسن تصميم الدليل الخاص ببيان ما على أساس تصنيف حالات هذا البيان أكثر تفصيلاً من التصنيف الذي سوف يظهر في التبويب النهائي ، حيث لا يمكن الحكم على مدى اتساع الفئات أو ضيقها الا بعد ظهور الجداول الأولى . فإذا تبين لنا من هذه الجداول المعدة على أساس التصنيف المفصل ان الفئات المستخدمة واسعة دون مبرر نظراً لوجود عدد قليل من الحالات في كل منها فإنه من السهل تجميع الفئات لإظهارها في الجداول النهائية. والعكس إذا تبين ازدحام الفئات بالتكرار فإنه يكون من الصعب تقسيم الفئات ويتحتم علينا في هذه الحالة إعادة العمل مرة أخرى على أساس أكثر تفصيلاً . لهذا وحتى نضمن لأنفسنا خط الرجعة يحسن التصنيف على أساس تفصيلي إلى أقصى درجة ممكنة .

٢ - إذا كان التبويب سوف يجري يدوياً يحسن الابتعاد عن الرموز الرقمية حيث من السهل الخطأ فيها . وفي هذه الحالة يمكن أن تكون الرموز عبارة عن حروف أبجدية لها علاقة بالكلمات التي تمثلها أو اختصارات لهذه الكلمات واستخدام هذه الطريقة يجنبنا كثيراً من الخطأ . وإذا أردنا استخدام الرموز كدليل يحسن تجنب الرموز التي يمكن ان تختلط فيها بينها عند كتابتها.

٣ - إذا كانت بعض البيانات عبارة عن أرقام (عدد العمال مثلاً) لا نحتاج الى توقيع دليل لها ويحسن إحاطة الرقم بدائرة بلون خاص حتى يتنبه المثقب اليها .

٤ - يجب مراجعة الترميز قبل البدء في العمليات الآلية الأخرى (التثقيب والفرز والتبويب) .

٥ - إذا كان البيان يحتاج لإجراء عمليات حسابية قبل توقيع الدليل الخاص به يحسن الفصل بين العمليتين فيقوم موظف بإجراء العمل الحسابي ومراجعته أولاً ثم يقوم موظف آخر بتوقيع الدليل .

٦ - يحسن أن يكون الرمزين ملين بالتعاريف والاصطلاحات المستخدمة في الدراسة ولذا يكون من الافضل توظيفهم من العدادين أو المراجعين الذين سبق لهم العمل في هذه الدراسة او من موظفي المكتب الذين لهم دراية بهذه التعاريف والمصطلحات .

٧ - يجب أن ينظم العمل في الترميز بحيث يجري العمل روتينياً ، فإذا وجدت مشاكل خاصة بالعمل يحسن ان يكون هناك موظف مختص لمعالجتها .

٨ - إذا احتوت استمارة البحث على بيانات صعبة الترميز (ترميز المهن أو النشاط الاقتصادي مثلاً) يكون من الافضل تقسيم عملية الترميز بحيث تعطى البيانات الصعبة لموظفين متخصصين في ترميزها وتعطى البيانات الأخرى لمرمزين آخرين ، وفي هذه الحالة يحسن ان يقوم بمراجعة الترميز ككل موظف واحد .

٩ - يجب أن تتم عملية الترميز بناء على تعليمات موجهة للمرمزين حتى يمكن توحيد خطوات العمل وطريقته بين جميع المرمزين القائمين بالعمل .

١٠ - ان دقة الترميز تتوقف أساساً على الدقة التي جمعت بها البيانات في استمارة البحث وعلى التوضيح الذي تسجل به تلك البيانات ، خصوصاً بالنسبة لبيانات

كذلك الخاصة بالمهنة او بالنشاط الاقتصادي. فإذا كنا بصدد النشاط الاقتصادي مثلاً يجب ان يكون البيان المسجل في الاستمارة موضعاً الى أبعد درجة حتى لا يختلط الأمر على المرمزين ، فلو وردت في الاستمارة عبارة « ملابس » لما كان ذلك كافياً لتوقيع الدليل الصحيح اذ يحتمل ان تكون المؤسسة مصنعة للملابس أو تجارة ملابس بالجملة أو تجارة ملابس بالفرق . وكل من تلك الاحتمالات يرشحها لدليل مختلف كل الاختلاف عن الآخر .

١١ - يجب أن يكون واضحاً للقائمين بالعمل في الترميز مدى الخطورة التي تترتب على عدم الدقة في ترميز البيانات حيث يكون ذلك بمثابة تغيير للبيانات ، الامر الذي يجعل النتائج غير دقيقة ومضللة . وواضح ان معنى ذلك القضاء على كل الجهود التي سوف تبذل عقب توقيع الدليل حتى ظهور النتائج .

اجراءات تنظيمية لعمليات الترميز :

يجب ان تكون الاجراءات اللازمة لتنظيم عملية الترميز قد أعدت وجهزت قبل البدء في العمل ، والاجراءات اللازمة هي :

١ - إعداد المطبوعات اللازمة للعملية - يجب إعداد نسخاً من الدليل الذي سوف يستخدم في العملية ، وتوزيعه على المرمزين ، ويحسن ان يشتمل الدليل المطبوع على مقدمة تشرح بنود الدليل المختلفة واذا كانت التصنيفات الدولية سوف تستخدم في الترميز (للمهن او النشاط الاقتصادي .. الخ) فيجب إعداد نسخاً منها مع مقدمة تشرح كيفية استعمالها في الترميز، ويحسن أن تحتوي هذه النسخ على فهرس حسب الحروف الهجائية لمساعدة المرمزين في العمل نظراً لتعقيد هذه التصنيفات وطولها . وأثناء العمل إذا أجريت بعض التعديلات أو إذا ظهرت بعض المشاكل التي أثارها المرمزين وقدمت لها الحلول يحسن إعداد مطبوعات بذلك وتوزيعها على الموظفين أولاً بأول . ويا حبذا

لو كان هناك تنظيم يمكن بمقتضاه أن يحتفظ كل مرمز بالمطبوعات الخاصة بالعمل في ملف خاص به .

٢ - اختيار الموظفين لعملية توقيع الدليل - يجب اختيار الموظفين الصالحين لهذه العملية وقد سبق أن أشرت الى انه يحسن ان يكونوا ملينين بالتعاريف والمصطلحات وكل الاسس التي يجري عليها العمل . وبالإضافة الى ذلك يجب ان يتوفر في الموظفين الذين يتم اختيارهم للترميز بعض الصفات مثل قوة الذاكرة وعدم النسيان والقدرة على التصرف وخاصة بالنسبة للموظفين الذين سوف يقوموا بالعمل مستخدمين أنواع الدليل المعقدة ، وكذلك يجب ان يتحلوا بالدقة وعدم الاستهتار . وتجري المفاضلة دائما بين اي الجنسين - الذكور او الاناث - اصلح لهذه العملية ويفضل البعض الاناث لتحليهن بالصبر وقدرتهن على الاعمال الرتيبة المكررة ، الا ان الذكور اثبتوا كذلك صلاحية في هذا العمل وخاصة بالنسبة للتصنيفات المعقدة التي تحتاج الى نوع من المرونة . لذلك عند بحث هذا الموضوع يجب أن لا يكون على أساس مقدرة كل من الجنسين وإنما على أساس ظروف كل بيئة وتقاليدها وتوفر العدد اللازم من الجنسين وحاجة كل منها للعمل .

٣ - تدريب الموظفين على توقيع الدليل - يجب تدريب الموظفين على العمل الذي سوف يوكل اليهم ، ويبدأ التدريب نظريا باعطاء الدروس لشرح الدليل وطريقة توقيعه وكيفية حفظه بسهولة . ومدة هذا التدريب تختلف تبعا لنوع الدليل المطلوب التدريب عليه من ناحية التعقيد او البساطة . وبعد التدريب النظري يبدأ التدريب العملي بتوزيع استمارات بها البيانات التي سوف ترمز على الموظفين ومطابقتها بترميزها تحت اشراف المدربين . وبالطبع لا يجب استخدام الاستمارات الاصلية حيث ان الموظفين سوف يرتكبون أخطاء كثيرة عندما يبدأون في العمل ولذلك يجب أعداد كشوف لهذا

الغرض ويكفي أن تحتوي على البيانات موضوع الترميز فقط . وعلى ضوء نتائج التدريب يمكن استبعاد من لا يصلحون لهذا العمل .

التعليمات الخاصة بالترميز :

سبق أن أشرت الى أن عملية الترميز يجب أن تجري بناء على تعليمات مفصلة وواضحة لتوحيد خطوات للعمل بين الموظفين . ويحسن أن تشمل هذه التعليمات على الآتي : -

١ - شرح الغرض من الترميز وموقع عملية الترميز من العملية الاحصائية بكاملها .

٢ - بيان بضرورة اطلاع المرمزين على التعليمات الموجهة للمعدادين والمراجعين والعاملين على الآلات الثاقبة والفارزة (ولهذا يحسن توزيع نسخاً من هذه التعليمات على المرمزين كذلك) .

٣ - الترميز يكون للبيانات بعد تصحيحها سواء أجرى التصحيح من قبل العداد أو المشرف أو المراجع . أي أن لا يجب الالتفات الى البيانات الأصلية الخاطئة .

٤ - تحديد بيانات الاستمارة التي سيوقع لها الدليل .

٥ - تحديد أنواع الدليل التي سوف تستخدم في العمل (مع ملاحظة ضرورة توفر نسخة من كل دليل يجري استخدامه مع كل موظف) .

٦ - شرح لكيفية تسجيل ارقام الدليل ، وهذا يشتمل على بيان المكان الذي يوقع فيه الدليل هل على يسار أم على يمين البيان ، أم فوقه أم تحته أم في خانة مخصصة لذلك . كذلك يجب تحديد لون القلم الذي يستخدم في توقيع الدليل (وعادة يستخدم اللون الازرق) ولون القلم الذي يستخدم في مراجعته الترميز (وعادة يستخدم اللون الاحمر) . كذلك يجب التنبيه على

الكتابة بخط واضح مقروء واذا حدث خطأ يشطب الرقم الخاطئ ويكتب فوقه الرقم المصحح ، أي لا يجب مسح رقم الدليل او تغييره عند حدوث خطأ . كذلك يجب الاشارة الى البيانات التي لن يجري ترميزها وانما يجب احاطتها بدائرة لتنبيه المثقبن اليها .

٧ - تحديد ترتيب العمل بالنسبة لترميز البيانات المختلفة التي تحتوي عليها الاستمارة . ويدور نقاش حول الطريقة التي يقوم بها الموظف بتوقيع الدليل ، هل يقوم بترميز بيان ما بالنسبة لجميع الوحدات ثم ينتقل الى بيان آخر أم يقوم بترميز جميع البيانات الخاصة بوحدة ما ثم ينتقل الى ترميز بيانات وحدة أخرى .

ان كلا من هاتين الطريقتين لها مزاياها وعيوبها . ففي الحالة الاولى يقل الجهد الذي يبذله الموظف حيث انه يركز ذهنه في نوع من واحد من البيانات فيعود عليه ويصبح من السهل توقيع رقم الدليل دون الحاجة الى البحث عنه طويلاً . الا أن الموظف لا يستطيع الربط بين بيانات الوحدة الواحدة ، لذلك لا تساعد هذه الطريقة على اكتشاف الأخطاء التي يمكن أن توجد في البيانات بالرغم من مراجعتها قبل ذلك .

أما بتطبيق الطريقة الثانية يمكن الربط بين بيانات الوحدة الواحدة ويستطيع الموظف تبعاً لذلك الاهتمام الى الدليل الصحيح لبيان ما قد لا يمكن الاهتمام اليه دوو هذا الربط . كذلك فان الربط بين البيانات يساعد في اكتشاف الأخطاء التي يحتمل وجودها فيها . إلا انه في هذه الحالة يبذل الموظف مجهوداً كبيراً يتمثل في الانتقال من بيان الى آخر بسرعة الأمر الذي لا يساعده كثيراً في حفظ أرقام الدليل .

والفصل في هذا النقاش يتوقف على مدى دقة عملية المراجعة ، فإذا كانت المراجعة قد اجريت بدقة ولنا ثقة كبيرة في القائمين بها فإن الطريقة الاولى تكون أفضل حيث انها تسهل العمل كثيراً على الموظفين والعكس صحيح

٨ - توجيه نظر الموظفين الى ضرورة الربط بين مختلف البيانات الخاصة بوحدة واحدة اذا كان هناك غموض يكتنف بعض البيانات ، فقد يؤدي الربط بين البيانات الى كشف هذا الغموض .

٩ - توجيه نظر الموظفين الى الحالات التي يوقع فيها دليل « لم يبين » وهي الحالات التي يعتبرها الغموض ويجب محاولة الوصول الى قبيانها باتباع ما ورد في الفقرة السابقة قبل ان يقرر الموظفون اعطاءها دليل « لم يبين » ، ويحسن في مثل هذه الحالات استشارة المشرف على عملية الترميز .

١٠ - لفت نظر الموظفين الى الاخطاء التي يمكن ان يقعوا فيها اذا لم ينتبهوا جيداً أثناء عملهم .

١١ - لفت نظر الموظفين الى الاخطاء التي يحتمل وجودها في البيانات والتي يجب عليهم تصحيحها قبل الاستمرار في عملية الترميز .

١٢ - تحديد الطريقة التي يمكن بواسطتها ان يستفسر الموظفون عن الصعوبات والمشاكل التي تواجههم في العمل على ان يكون التبليغ عنها كتابة .

١٣ - توجيه نظر الموظفين الى متابعة التعليقات التي تظهر أولاً بأول وان يحتفظوا بنسخة منها لديهم .

١٤ - تحديد المسائل التنظيمية المختلفة مثل : -

أ - كيفية استلام الاستمارات .

ب - كيفية اعادتها بعد ترميزها .

ج - كيفية اعداد التقارير عن سير العمل بالنسبة لكل موظف .

د - مواعيد تقديم هذه التقارير .

هـ - المعدل اليومي لانتاج الموظفين .

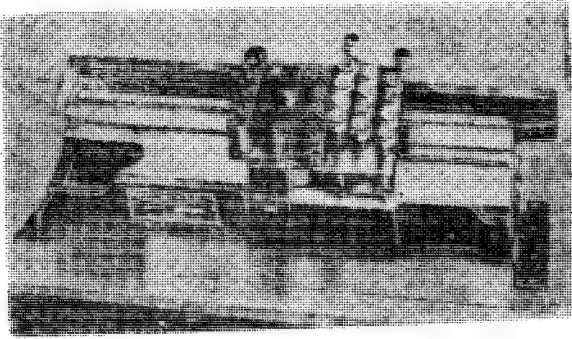
١٥ - تقديم مثال لبيانات على استمارة و ترميزها بكل دقة وعناية .

١٦ - النص على أن يوقع المرمزون ومراجعو الترميز على الاستمارات التي ينتهون منها وان يكون التوقيع مصحوباً بالتاريخ ومكان ذلك التوقيع .

١٧ - النص على تقسيم العمل بين الموظفين اذا كان هناك رغبة في ذلك ، حيث يفضل البعض أن يقوم بعض الموظفين بترميز البيانات الصعبة مثل المهنة والنشاط الاقتصادي بينما يقوم الموظفون الباقون بترميز البيانات الاخرى . وعند اتباع هذه الطريقة يراعى في الفريق الذي يقوم بتوقيع الدليل الصعب أن يكون أكثر مرونة ودقة في العمل . على أن اتباع هذه الطريقة يؤدي الى تضخيم حجم العمل والقسم القائم به الى حد ما الا انه في نفس الوقت ادعى الى الدقة والسلامة اذا أمكن التغلب على العيب الناشيء من ضخامة العمل

التثقيب ومراجعته :

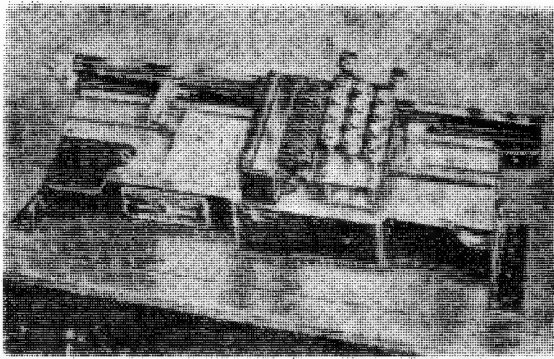
والخطوة التالية هي تحويل الرموز التي كتبت أمام الاجابات المختلفة إلى ثقب في مواضع معينة من بطاقة التثقيب . ولإجراء ذلك يخصص لكل سؤال (لكل بيان) عمود أو أكثر في البطاقة بحسب عدد الأرقام التي يتكون منها آخر عدد في دليل هذا البيان . فإذا كان دليل البيان مكون من رقم واحد ، أي إذا كانت الاجابات المحتملة للسؤال لا تصل الى العدد ١٠ خصص للسؤال عموداً واحداً في البطاقة ، وإذا كان دليل اجابات السؤال يتكون من رقمين أي أن الاجابات المحتملة لهذا السؤال يمكن ان تصل الى ٩٩ في الحد الأقصى خصصنا للسؤال عمودين احدهما لرقم الآحاد والثاني لرقم العشرات . وتجري عملية التثقيب بعمل ثقب (آلياً) في الأعمدة المخصصة لاجابات السؤال وأمام الأرقام الخاصة باجابة معينة في الاستارة ، مثلاً اذا كان رقم الدليل لاجابة معينة هو ٢٥ يعمل ثقب أمام رقم ٢ في عمود العشرات وثقب أمام رقم ٥ في عمود الآحاد .



آلة التثقيب اليدوية

وتجري عملية التثقيب بواسطة آلة تسمى بآلة التثقيب وهي تحتوي على عشرة أزرار مكتوب على كل منها أحد الأرقام من صفر الى ٩ . وعند الضغط على أحد الأزرار وليكن الزر رقم ٨ تقوم الآلة بعمل ثقب في مكان الرقم ٨ في البطاقة . وعند بدء العمل توضع البطاقة في داخل الآلة ، وبذلك تكون الآلة مستعدة للتثقيب في العمود الاول من البطاقة ، فإذا ضغطنا على الزر الخاص برقم معين فتثقبه الآلة في هذا العمود وتنتقل البطاقة آلياً الى العمود الثاني أي تندفع البطاقة آلياً إلى الخارج مسافة مساوية للمسافة الواقعة بين عمودين من أعمدها (وهذه الحركة مماثلة لحركة قطعة الورق في حالة الآلات الكاتبة) ، ويستمر الثاقب في الضغط على الأزرار حسب ارقام الدليل المدونة امام الاجابات المختلفة حسب ترتيبها في اعمدة بطاقة التثقيب حتى ينتهي منها جميعاً . وعندئذ يضغط على زر خاص فتخرج البطاقة ويسحبها الثاقب خارج الآلة ثم يضع بطاقة ثانية للوحدة الثانية ويشرح في تثقيبها وهكذا . وبذلك نلاحظ ان كمية البيانات التي نثقبها في بطاقة التثيب تتوقف على كمية البيانات التي نريد تثقيبها وعلى عدد الاعمدة التي يحتاجها دليل هذه البيانات وعلى عدد

الاعمدة الموجود فعلا في بطاقة التثقيب ، فاذا كانت البيانات الخاصة بكل وحدة من وحدات الدراسة يحتاج تثقيبها الى ٧٩ عمود وكانت اعمدة البطاقة هي ٦٥ فقط نحتاج بذلك الى بطاقتين لتثقيب البيانات الخاصة بوحدة واحدة ، اما اذا كانت اعمدة البطاقة ٨٥ عمودا نحتاج بذلك الى بطاقة واحدة لبيانات كل وحدة . لهذا نلاحظ ان كثرة اعمدة البطاقة تعتبر ميزة كبيرة في بطاقات التثقيب ، والموجود في الوقت الحاضر من هذه البطاقات يحتوي اما على ٣٦ عمود (وهو صنف قديم جداً) لم يعد يستعمل كثيرا ، أو ٦٥ عمود أو ٨٥ عمود . وبالطبع كل نوع من آلات الاحصائية يكون معدا له نوع معين من بطاقات التثقيب (من حيث عدد الاعمدة) ، ولذلك عندما نتكلم عن آلات الاحصائية نقول آلات ٣٦ عمود او آلات ٦٥ عمود او آلات ٨٥ عمود .



آلة المراجعة اليدوية

وحيث انه يمكن ان تحدث اخطاء عند اجراء عملية التثقيب ، اذ قد يضغط الثاقب على الزر الخاص برقم ٨ بينما يكون رقم الدليل ٧ ، لهذا

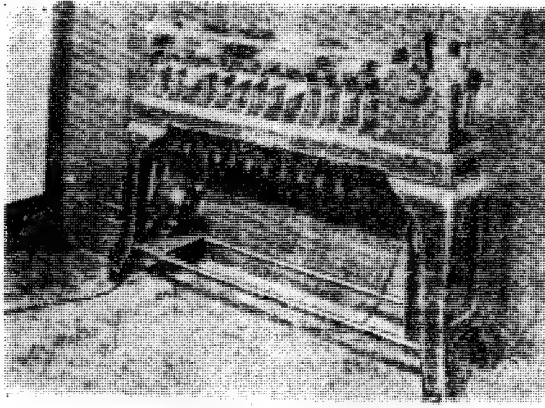
يكون من الواجب مراجعة هذه العملية . وتم عملية المراجعة آليا كذلك بواسطة آلة خاصة بذلك لا تختلف كثيرا عن آلة التثقيب السابقة . ويقوم مراجع التثقيب بادخال البطاقة المثقبة في آلة المراجعة ويضغط على الازرار وفقا لارقام الدليل الخاصة بالاجابات المختلفة وتبعا لترتيبها حسب اعمدة البطاقة ، فاذا كان الرقم الذي يضغط عليه يطابق الرقم المثقوب في البطاقة تندفع البطاقة الى الخارج ويدل ذلك على صحة التثقيب وبذلك يستمر المراجع في عمله ، أما اذا كان الرقم الذي يضغط عليه لا يطابق الرقم المثقوب في البطاقة لا تتحرك وحينئذ ينتبه المراجع الى وجود خطأ في التثقيب ، وحيث انه يمكن أن يكون هو الذي أخطأ لذلك يجب أن يتأكد من صحة ضغطه على الزر فاذا كان الأمر كذلك علم أن هناك خطأ في عمل المثقب وبذلك يسحب البطاقة من الآلة حتى تمزق وتعمل بطاقة اخرى بدلها .

التصنيف (الفرز) :

باجراء الخطوات السابقة نكون قد حولنا البيانات الموجودة في الاستمارة الى ثقب في أعمدة البطاقات طبقاً لارقام الدليل الخاصة بها ، وبذلك يتجمع لدينا بطاقات مثقبة بحيث يكون لكل وحدة استبيان بطاقة أو أكثر حسب نوع البطاقة التي نستخدمها وكمية المعلومات التي ثقبناها لكل وحدة ، والأجراء التالي هو تصنيف هذه البطاقات أي فرزها بالنسبة لكل بيان من البيانات التي ثقبناها فاذا كان لدينا بطاقات تمثل المعلومات المختلفة عن السكان وأردنا مثلاً تصنيف السكان تبعاً للجنس ، لذلك نعرف أولاً العمود المخصص لتثقيب هذا البيان وليكن العمود رقم ٥ ، ثم نمرر البطاقات في الات كهربائية تسمى بالآلات الفرز حيث تضبط مؤشر الآلة على العمود رقم ٥ (يوجد على الآلة مقياس مدرج مكتوب عليه أرقام الأعمدة الموجودة في البطاقة ، أي الأرقام ١، ٢، ٣، ٤، ٥، ٦، ٧، ٨، ٩، ١٠، ١١، ١٢، ١٣، ١٤، ١٥، ١٦، ١٧، ١٨، ١٩، ٢٠، ٢١، ٢٢، ٢٣، ٢٤، ٢٥، ٢٦، ٢٧، ٢٨، ٢٩، ٣٠، ٣١، ٣٢، ٣٣، ٣٤، ٣٥، ٣٦، ٣٧، ٣٨، ٣٩، ٤٠، ٤١، ٤٢، ٤٣، ٤٤، ٤٥، ٤٦، ٤٧، ٤٨، ٤٩، ٥٠، ٥١، ٥٢، ٥٣، ٥٤، ٥٥، ٥٦، ٥٧، ٥٨، ٥٩، ٦٠، ٦١، ٦٢، ٦٣، ٦٤، ٦٥، ٦٦، ٦٧، ٦٨، ٦٩، ٧٠، ٧١، ٧٢، ٧٣، ٧٤، ٧٥، ٧٦، ٧٧، ٧٨، ٧٩، ٨٠، ٨١، ٨٢، ٨٣، ٨٤، ٨٥، ٨٦، ٨٧، ٨٨، ٨٩، ٩٠، ٩١، ٩٢، ٩٣، ٩٤، ٩٥، ٩٦، ٩٧، ٩٨، ٩٩، ١٠٠، ١٠١، ١٠٢، ١٠٣، ١٠٤، ١٠٥، ١٠٦، ١٠٧، ١٠٨، ١٠٩، ١١٠، ١١١، ١١٢، ١١٣، ١١٤، ١١٥، ١١٦، ١١٧، ١١٨، ١١٩، ١٢٠، ١٢١، ١٢٢، ١٢٣، ١٢٤، ١٢٥، ١٢٦، ١٢٧، ١٢٨، ١٢٩، ١٣٠، ١٣١، ١٣٢، ١٣٣، ١٣٤، ١٣٥، ١٣٦، ١٣٧، ١٣٨، ١٣٩، ١٤٠، ١٤١، ١٤٢، ١٤٣، ١٤٤، ١٤٥، ١٤٦، ١٤٧، ١٤٨، ١٤٩، ١٥٠، ١٥١، ١٥٢، ١٥٣، ١٥٤، ١٥٥، ١٥٦، ١٥٧، ١٥٨، ١٥٩، ١٦٠، ١٦١، ١٦٢، ١٦٣، ١٦٤، ١٦٥، ١٦٦، ١٦٧، ١٦٨، ١٦٩، ١٧٠، ١٧١، ١٧٢، ١٧٣، ١٧٤، ١٧٥، ١٧٦، ١٧٧، ١٧٨، ١٧٩، ١٨٠، ١٨١، ١٨٢، ١٨٣، ١٨٤، ١٨٥، ١٨٦، ١٨٧، ١٨٨، ١٨٩، ١٩٠، ١٩١، ١٩٢، ١٩٣، ١٩٤، ١٩٥، ١٩٦، ١٩٧، ١٩٨، ١٩٩، ٢٠٠، ٢٠١، ٢٠٢، ٢٠٣، ٢٠٤، ٢٠٥، ٢٠٦، ٢٠٧، ٢٠٨، ٢٠٩، ٢١٠، ٢١١، ٢١٢، ٢١٣، ٢١٤، ٢١٥، ٢١٦، ٢١٧، ٢١٨، ٢١٩، ٢٢٠، ٢٢١، ٢٢٢، ٢٢٣، ٢٢٤، ٢٢٥، ٢٢٦، ٢٢٧، ٢٢٨، ٢٢٩، ٢٣٠، ٢٣١، ٢٣٢، ٢٣٣، ٢٣٤، ٢٣٥، ٢٣٦، ٢٣٧، ٢٣٨، ٢٣٩، ٢٤٠، ٢٤١، ٢٤٢، ٢٤٣، ٢٤٤، ٢٤٥، ٢٤٦، ٢٤٧، ٢٤٨، ٢٤٩، ٢٥٠، ٢٥١، ٢٥٢، ٢٥٣، ٢٥٤، ٢٥٥، ٢٥٦، ٢٥٧، ٢٥٨، ٢٥٩، ٢٦٠، ٢٦١، ٢٦٢، ٢٦٣، ٢٦٤، ٢٦٥، ٢٦٦، ٢٦٧، ٢٦٨، ٢٦٩، ٢٧٠، ٢٧١، ٢٧٢، ٢٧٣، ٢٧٤، ٢٧٥، ٢٧٦، ٢٧٧، ٢٧٨، ٢٧٩، ٢٨٠، ٢٨١، ٢٨٢، ٢٨٣، ٢٨٤، ٢٨٥، ٢٨٦، ٢٨٧، ٢٨٨، ٢٨٩، ٢٩٠، ٢٩١، ٢٩٢، ٢٩٣، ٢٩٤، ٢٩٥، ٢٩٦، ٢٩٧، ٢٩٨، ٢٩٩، ٣٠٠، ٣٠١، ٣٠٢، ٣٠٣، ٣٠٤، ٣٠٥، ٣٠٦، ٣٠٧، ٣٠٨، ٣٠٩، ٣١٠، ٣١١، ٣١٢، ٣١٣، ٣١٤، ٣١٥، ٣١٦، ٣١٧، ٣١٨، ٣١٩، ٣٢٠، ٣٢١، ٣٢٢، ٣٢٣، ٣٢٤، ٣٢٥، ٣٢٦، ٣٢٧، ٣٢٨، ٣٢٩، ٣٣٠، ٣٣١، ٣٣٢، ٣٣٣، ٣٣٤، ٣٣٥، ٣٣٦، ٣٣٧، ٣٣٨، ٣٣٩، ٣٤٠، ٣٤١، ٣٤٢، ٣٤٣، ٣٤٤، ٣٤٥، ٣٤٦، ٣٤٧، ٣٤٨، ٣٤٩، ٣٥٠، ٣٥١، ٣٥٢، ٣٥٣، ٣٥٤، ٣٥٥، ٣٥٦، ٣٥٧، ٣٥٨، ٣٥٩، ٣٦٠، ٣٦١، ٣٦٢، ٣٦٣، ٣٦٤، ٣٦٥، ٣٦٦، ٣٦٧، ٣٦٨، ٣٦٩، ٣٧٠، ٣٧١، ٣٧٢، ٣٧٣، ٣٧٤، ٣٧٥، ٣٧٦، ٣٧٧، ٣٧٨، ٣٧٩، ٣٨٠، ٣٨١، ٣٨٢، ٣٨٣، ٣٨٤، ٣٨٥، ٣٨٦، ٣٨٧، ٣٨٨، ٣٨٩، ٣٩٠، ٣٩١، ٣٩٢، ٣٩٣، ٣٩٤، ٣٩٥، ٣٩٦، ٣٩٧، ٣٩٨، ٣٩٩، ٤٠٠، ٤٠١، ٤٠٢، ٤٠٣، ٤٠٤، ٤٠٥، ٤٠٦، ٤٠٧، ٤٠٨، ٤٠٩، ٤١٠، ٤١١، ٤١٢، ٤١٣، ٤١٤، ٤١٥، ٤١٦، ٤١٧، ٤١٨، ٤١٩، ٤٢٠، ٤٢١، ٤٢٢، ٤٢٣، ٤٢٤، ٤٢٥، ٤٢٦، ٤٢٧، ٤٢٨، ٤٢٩، ٤٣٠، ٤٣١، ٤٣٢، ٤٣٣، ٤٣٤، ٤٣٥، ٤٣٦، ٤٣٧، ٤٣٨، ٤٣٩، ٤٤٠، ٤٤١، ٤٤٢، ٤٤٣، ٤٤٤، ٤٤٥، ٤٤٦، ٤٤٧، ٤٤٨، ٤٤٩، ٤٥٠، ٤٥١، ٤٥٢، ٤٥٣، ٤٥٤، ٤٥٥، ٤٥٦، ٤٥٧، ٤٥٨، ٤٥٩، ٤٦٠، ٤٦١، ٤٦٢، ٤٦٣، ٤٦٤، ٤٦٥، ٤٦٦، ٤٦٧، ٤٦٨، ٤٦٩، ٤٧٠، ٤٧١، ٤٧٢، ٤٧٣، ٤٧٤، ٤٧٥، ٤٧٦، ٤٧٧، ٤٧٨، ٤٧٩، ٤٨٠، ٤٨١، ٤٨٢، ٤٨٣، ٤٨٤، ٤٨٥، ٤٨٦، ٤٨٧، ٤٨٨، ٤٨٩، ٤٩٠، ٤٩١، ٤٩٢، ٤٩٣، ٤٩٤، ٤٩٥، ٤٩٦، ٤٩٧، ٤٩٨، ٤٩٩، ٥٠٠، ٥٠١، ٥٠٢، ٥٠٣، ٥٠٤، ٥٠٥، ٥٠٦، ٥٠٧، ٥٠٨، ٥٠٩، ٥١٠، ٥١١، ٥١٢، ٥١٣، ٥١٤، ٥١٥، ٥١٦، ٥١٧، ٥١٨، ٥١٩، ٥٢٠، ٥٢١، ٥٢٢، ٥٢٣، ٥٢٤، ٥٢٥، ٥٢٦، ٥٢٧، ٥٢٨، ٥٢٩، ٥٣٠، ٥٣١، ٥٣٢، ٥٣٣، ٥٣٤، ٥٣٥، ٥٣٦، ٥٣٧، ٥٣٨، ٥٣٩، ٥٤٠، ٥٤١، ٥٤٢، ٥٤٣، ٥٤٤، ٥٤٥، ٥٤٦، ٥٤٧، ٥٤٨، ٥٤٩، ٥٥٠، ٥٥١، ٥٥٢، ٥٥٣، ٥٥٤، ٥٥٥، ٥٥٦، ٥٥٧، ٥٥٨، ٥٥٩، ٥٦٠، ٥٦١، ٥٦٢، ٥٦٣، ٥٦٤، ٥٦٥، ٥٦٦، ٥٦٧، ٥٦٨، ٥٦٩، ٥٧٠، ٥٧١، ٥٧٢، ٥٧٣، ٥٧٤، ٥٧٥، ٥٧٦، ٥٧٧، ٥٧٨، ٥٧٩، ٥٨٠، ٥٨١، ٥٨٢، ٥٨٣، ٥٨٤، ٥٨٥، ٥٨٦، ٥٨٧، ٥٨٨، ٥٨٩، ٥٩٠، ٥٩١، ٥٩٢، ٥٩٣، ٥٩٤، ٥٩٥، ٥٩٦، ٥٩٧، ٥٩٨، ٥٩٩، ٦٠٠، ٦٠١، ٦٠٢، ٦٠٣، ٦٠٤، ٦٠٥، ٦٠٦، ٦٠٧، ٦٠٨، ٦٠٩، ٦١٠، ٦١١، ٦١٢، ٦١٣، ٦١٤، ٦١٥، ٦١٦، ٦١٧، ٦١٨، ٦١٩، ٦٢٠، ٦٢١، ٦٢٢، ٦٢٣، ٦٢٤، ٦٢٥، ٦٢٦، ٦٢٧، ٦٢٨، ٦٢٩، ٦٣٠، ٦٣١، ٦٣٢، ٦٣٣، ٦٣٤، ٦٣٥، ٦٣٦، ٦٣٧، ٦٣٨، ٦٣٩، ٦٤٠، ٦٤١، ٦٤٢، ٦٤٣، ٦٤٤، ٦٤٥، ٦٤٦، ٦٤٧، ٦٤٨، ٦٤٩، ٦٥٠، ٦٥١، ٦٥٢، ٦٥٣، ٦٥٤، ٦٥٥، ٦٥٦، ٦٥٧، ٦٥٨، ٦٥٩، ٦٦٠، ٦٦١، ٦٦٢، ٦٦٣، ٦٦٤، ٦٦٥، ٦٦٦، ٦٦٧، ٦٦٨، ٦٦٩، ٦٧٠، ٦٧١، ٦٧٢، ٦٧٣، ٦٧٤، ٦٧٥، ٦٧٦، ٦٧٧، ٦٧٨، ٦٧٩، ٦٨٠، ٦٨١، ٦٨٢، ٦٨٣، ٦٨٤، ٦٨٥، ٦٨٦، ٦٨٧، ٦٨٨، ٦٨٩، ٦٩٠، ٦٩١، ٦٩٢، ٦٩٣، ٦٩٤، ٦٩٥، ٦٩٦، ٦٩٧، ٦٩٨، ٦٩٩، ٧٠٠، ٧٠١، ٧٠٢، ٧٠٣، ٧٠٤، ٧٠٥، ٧٠٦، ٧٠٧، ٧٠٨، ٧٠٩، ٧١٠، ٧١١، ٧١٢، ٧١٣، ٧١٤، ٧١٥، ٧١٦، ٧١٧، ٧١٨، ٧١٩، ٧٢٠، ٧٢١، ٧٢٢، ٧٢٣، ٧٢٤، ٧٢٥، ٧٢٦، ٧٢٧، ٧٢٨، ٧٢٩، ٧٣٠، ٧٣١، ٧٣٢، ٧٣٣، ٧٣٤، ٧٣٥، ٧٣٦، ٧٣٧، ٧٣٨، ٧٣٩، ٧٤٠، ٧٤١، ٧٤٢، ٧٤٣، ٧٤٤، ٧٤٥، ٧٤٦، ٧٤٧، ٧٤٨، ٧٤٩، ٧٥٠، ٧٥١، ٧٥٢، ٧٥٣، ٧٥٤، ٧٥٥، ٧٥٦، ٧٥٧، ٧٥٨، ٧٥٩، ٧٦٠، ٧٦١، ٧٦٢، ٧٦٣، ٧٦٤، ٧٦٥، ٧٦٦، ٧٦٧، ٧٦٨، ٧٦٩، ٧٧٠، ٧٧١، ٧٧٢، ٧٧٣، ٧٧٤، ٧٧٥، ٧٧٦، ٧٧٧، ٧٧٨، ٧٧٩، ٧٨٠، ٧٨١، ٧٨٢، ٧٨٣، ٧٨٤، ٧٨٥، ٧٨٦، ٧٨٧، ٧٨٨، ٧٨٩، ٧٩٠، ٧٩١، ٧٩٢، ٧٩٣، ٧٩٤، ٧٩٥، ٧٩٦، ٧٩٧، ٧٩٨، ٧٩٩، ٨٠٠، ٨٠١، ٨٠٢، ٨٠٣، ٨٠٤، ٨٠٥، ٨٠٦، ٨٠٧، ٨٠٨، ٨٠٩، ٨١٠، ٨١١، ٨١٢، ٨١٣، ٨١٤، ٨١٥، ٨١٦، ٨١٧، ٨١٨، ٨١٩، ٨٢٠، ٨٢١، ٨٢٢، ٨٢٣، ٨٢٤، ٨٢٥، ٨٢٦، ٨٢٧، ٨٢٨، ٨٢٩، ٨٣٠، ٨٣١، ٨٣٢، ٨٣٣، ٨٣٤، ٨٣٥، ٨٣٦، ٨٣٧، ٨٣٨، ٨٣٩، ٨٤٠، ٨٤١، ٨٤٢، ٨٤٣، ٨٤٤، ٨٤٥، ٨٤٦، ٨٤٧، ٨٤٨، ٨٤٩، ٨٥٠، ٨٥١، ٨٥٢، ٨٥٣، ٨٥٤، ٨٥٥، ٨٥٦، ٨٥٧، ٨٥٨، ٨٥٩، ٨٦٠، ٨٦١، ٨٦٢، ٨٦٣، ٨٦٤، ٨٦٥، ٨٦٦، ٨٦٧، ٨٦٨، ٨٦٩، ٨٧٠، ٨٧١، ٨٧٢، ٨٧٣، ٨٧٤، ٨٧٥، ٨٧٦، ٨٧٧، ٨٧٨، ٨٧٩، ٨٨٠، ٨٨١، ٨٨٢، ٨٨٣، ٨٨٤، ٨٨٥، ٨٨٦، ٨٨٧، ٨٨٨، ٨٨٩، ٨٩٠، ٨٩١، ٨٩٢، ٨٩٣، ٨٩٤، ٨٩٥، ٨٩٦، ٨٩٧، ٨٩٨، ٨٩٩، ٩٠٠، ٩٠١، ٩٠٢، ٩٠٣، ٩٠٤، ٩٠٥، ٩٠٦، ٩٠٧، ٩٠٨، ٩٠٩، ٩١٠، ٩١١، ٩١٢، ٩١٣، ٩١٤، ٩١٥، ٩١٦، ٩١٧، ٩١٨، ٩١٩، ٩٢٠، ٩٢١، ٩٢٢، ٩٢٣، ٩٢٤، ٩٢٥، ٩٢٦، ٩٢٧، ٩٢٨، ٩٢٩، ٩٣٠، ٩٣١، ٩٣٢، ٩٣٣، ٩٣٤، ٩٣٥، ٩٣٦، ٩٣٧، ٩٣٨، ٩٣٩، ٩٤٠، ٩٤١، ٩٤٢، ٩٤٣، ٩٤٤، ٩٤٥، ٩٤٦، ٩٤٧، ٩٤٨، ٩٤٩، ٩٥٠، ٩٥١، ٩٥٢، ٩٥٣، ٩٥٤، ٩٥٥، ٩٥٦، ٩٥٧، ٩٥٨، ٩٥٩، ٩٦٠، ٩٦١، ٩٦٢، ٩٦٣، ٩٦٤، ٩٦٥، ٩٦٦، ٩٦٧، ٩٦٨، ٩٦٩، ٩٧٠، ٩٧١، ٩٧٢، ٩٧٣، ٩٧٤، ٩٧٥، ٩٧٦، ٩٧٧، ٩٧٨، ٩٧٩، ٩٨٠، ٩٨١، ٩٨٢، ٩٨٣، ٩٨٤، ٩٨٥، ٩٨٦، ٩٨٧، ٩٨٨، ٩٨٩، ٩٩٠، ٩٩١، ٩٩٢، ٩٩٣، ٩٩٤، ٩٩٥، ٩٩٦، ٩٩٧، ٩٩٨، ٩٩٩، ١٠٠٠، ١٠٠١، ١٠٠٢، ١٠٠٣، ١٠٠٤، ١٠٠٥، ١٠٠٦، ١٠٠٧، ١٠٠٨، ١٠٠٩، ١٠١٠، ١٠١١، ١٠١٢، ١٠١٣، ١٠١٤، ١٠١٥، ١٠١٦، ١٠١٧، ١٠١٨، ١٠١٩، ١٠٢٠، ١٠٢١، ١٠٢٢، ١٠٢٣، ١٠٢٤، ١٠٢٥، ١٠٢٦، ١٠٢٧، ١٠٢٨، ١٠٢٩، ١٠٣٠، ١٠٣١، ١٠٣٢، ١٠٣٣، ١٠٣٤، ١٠٣٥، ١٠٣٦، ١٠٣٧، ١٠٣٨، ١٠٣٩، ١٠٤٠، ١٠٤١، ١٠٤٢، ١٠٤٣، ١٠٤٤، ١٠٤٥، ١٠٤٦، ١٠٤٧، ١٠٤٨، ١٠٤٩، ١٠٥٠، ١٠٥١، ١٠٥٢، ١٠٥٣، ١٠٥٤، ١٠٥٥، ١٠٥٦، ١٠٥٧، ١٠٥٨، ١٠٥٩، ١٠٦٠، ١٠٦١، ١٠٦٢، ١٠٦٣، ١٠٦٤، ١٠٦٥، ١٠٦٦، ١٠٦٧، ١٠٦٨، ١٠٦٩، ١٠٧٠، ١٠٧١، ١٠٧٢، ١٠٧٣، ١٠٧٤، ١٠٧٥، ١٠٧٦، ١٠٧٧، ١٠٧٨، ١٠٧٩، ١٠٨٠، ١٠٨١، ١٠٨٢، ١٠٨٣، ١٠٨٤، ١٠٨٥، ١٠٨٦، ١٠٨٧، ١٠٨٨، ١٠٨٩، ١٠٩٠، ١٠٩١، ١٠٩٢، ١٠٩٣، ١٠٩٤، ١٠٩٥، ١٠٩٦، ١٠٩٧، ١٠٩٨، ١٠٩٩، ١١٠٠، ١١٠١، ١١٠٢، ١١٠٣، ١١٠٤، ١١٠٥، ١١٠٦، ١١٠٧، ١١٠٨، ١١٠٩، ١١١٠، ١١١١، ١١١٢، ١١١٣، ١١١٤، ١١١٥، ١١١٦، ١١١٧، ١١١٨، ١١١٩، ١١٢٠، ١١٢١، ١١٢٢، ١١٢٣، ١١٢٤، ١١٢٥، ١١٢٦، ١١٢٧، ١١٢٨، ١١٢٩، ١١٣٠، ١١٣١، ١١٣٢، ١١٣٣، ١١٣٤، ١١٣٥، ١١٣٦، ١١٣٧، ١١٣٨، ١١٣٩، ١١٤٠، ١١٤١، ١١٤٢، ١١٤٣، ١١٤٤، ١١٤٥، ١١٤٦، ١١٤٧، ١١٤٨، ١١٤٩، ١١٥٠، ١١٥١، ١١٥٢، ١١٥٣، ١١٥٤، ١١٥٥، ١١٥٦، ١١٥٧، ١١٥٨، ١١٥٩، ١١٦٠، ١١٦١، ١١٦٢، ١١٦٣، ١١٦٤، ١١٦٥، ١١٦٦، ١١٦٧، ١١٦٨، ١١٦٩، ١١٧٠، ١١٧١، ١١٧٢، ١١٧٣، ١١٧٤، ١١٧٥، ١١٧٦، ١١٧٧، ١١٧٨، ١١٧٩، ١١٨٠، ١١٨١، ١١٨٢، ١١٨٣، ١١٨٤، ١١٨٥، ١١٨٦، ١١٨٧، ١١٨٨، ١١٨٩، ١١٩٠، ١١٩١، ١١٩٢، ١١٩٣، ١١٩٤، ١١٩٥، ١١٩٦، ١١٩٧، ١١٩٨، ١١٩٩، ١٢٠٠، ١٢٠١، ١٢٠٢، ١٢٠٣، ١٢٠٤، ١٢٠٥، ١٢٠٦، ١٢٠٧، ١٢٠٨، ١٢٠٩، ١٢١٠، ١٢١١، ١٢١٢، ١٢١٣، ١٢١٤، ١٢١٥، ١٢١٦، ١٢١٧، ١٢١٨، ١٢١٩، ١٢٢٠، ١٢٢١، ١٢٢٢، ١٢٢٣، ١٢٢٤، ١٢٢٥، ١٢٢٦، ١٢٢٧، ١٢٢٨، ١٢٢٩، ١٢٣٠، ١٢٣١، ١٢٣٢، ١٢٣٣، ١٢٣٤، ١٢٣٥، ١٢٣٦، ١٢٣٧، ١٢٣٨، ١٢٣٩، ١٢٤٠، ١٢٤١، ١٢٤٢، ١٢٤٣، ١٢٤٤، ١٢٤٥، ١٢٤٦، ١٢٤٧، ١٢٤٨، ١٢٤٩، ١٢٥٠، ١٢٥١، ١٢٥٢، ١٢٥٣، ١٢٥٤، ١٢٥٥، ١٢٥٦، ١٢٥٧، ١٢٥٨، ١٢٥٩، ١٢٦٠، ١٢٦١، ١٢٦٢، ١٢٦٣، ١٢٦٤، ١٢٦٥، ١٢٦٦، ١٢٦٧، ١٢٦٨، ١٢٦٩، ١٢٧٠، ١٢٧١، ١٢٧٢، ١٢٧٣، ١٢٧٤، ١٢٧٥، ١٢٧٦، ١٢٧٧، ١٢٧٨، ١٢٧٩، ١٢٨٠، ١٢٨١، ١٢٨٢، ١٢٨٣، ١٢٨٤، ١٢٨٥، ١٢٨٦، ١٢٨٧، ١٢٨٨، ١٢٨٩، ١٢٩٠، ١٢٩١، ١٢٩٢، ١٢٩٣، ١٢٩٤، ١٢٩٥، ١٢٩٦، ١٢٩٧، ١٢٩٨، ١٢٩٩، ١٣٠٠، ١٣٠١، ١٣٠٢، ١٣٠٣، ١٣٠٤، ١٣٠٥، ١٣٠٦، ١٣٠٧، ١٣٠٨، ١٣٠٩، ١٣١٠، ١٣١١، ١٣١٢، ١٣١٣، ١٣١٤، ١٣١٥، ١٣١٦، ١٣١٧، ١٣١٨، ١٣١٩، ١٣٢٠، ١٣٢١، ١٣٢٢، ١٣٢٣، ١٣٢٤، ١٣٢٥، ١٣٢٦، ١٣٢٧، ١٣٢٨، ١٣٢٩، ١٣٣٠، ١٣٣١، ١٣٣٢، ١٣٣٣، ١٣٣٤، ١٣٣٥، ١٣٣٦، ١٣٣٧، ١٣٣٨، ١٣٣٩، ١٣٤٠، ١٣٤١، ١٣٤٢، ١٣٤٣، ١٣٤٤، ١٣٤٥، ١٣٤٦، ١٣٤٧، ١٣٤٨، ١٣٤٩، ١٣٥٠، ١٣٥١، ١٣٥٢، ١٣٥٣، ١٣٥٤، ١٣٥٥، ١٣٥٦، ١٣٥٧، ١٣٥٨، ١٣٥٩، ١٣٦٠، ١٣٦١، ١٣٦٢، ١٣٦٣، ١٣٦٤، ١٣٦٥، ١٣٦٦، ١٣٦٧، ١٣٦٨، ١٣٦٩، ١٣٧٠، ١٣٧١، ١٣٧٢، ١٣٧٣، ١٣٧٤، ١٣٧٥، ١٣٧٦، ١٣٧٧، ١٣٧٨، ١٣٧٩، ١٣٨٠، ١٣٨١، ١٣٨٢، ١٣٨٣، ١٣٨٤، ١٣٨٥، ١٣٨٦، ١٣٨٧، ١٣٨٨، ١٣٨٩، ١٣٩٠، ١٣٩١، ١٣٩٢، ١٣٩٣، ١٣٩٤، ١٣٩٥، ١٣٩٦، ١٣٩٧، ١٣٩٨، ١٣٩٩، ١٤٠٠، ١٤٠١، ١٤٠٢، ١٤٠٣، ١٤٠٤، ١٤٠٥، ١٤٠٦، ١٤٠٧، ١٤٠٨

المرجع فان ذلك يعني ان الآلة ستقوم بتصنيف أي فرز البطاقات حسب الأرقام المثقبة في العمود الذي يحمل هذا الرقم في البطاقة (. ولذلك بضبطنا المؤشر أمام رقم ٥ تكون الآلة مستعدة لفرز البطاقات حسب الأرقام المثقبة في هذا العمود فاذا كان دليل الجنس هو ١ للذكور و ٢ للإناث ، فان ادارتنا للآلة تجعلها تدفع البطاقات دفعا سريعا داخل مجرى معين وأثناء اندفاعها تسقط البطاقات المثقبة عند الرقم ١ في العمود الخامس فقط في جيب خاص بالرقم ١ ، أما البطاقات المثقبة عند الرقم ٢ في العمود الخامس فتسقط في جيب خاص بالرقم ٢ أما اذا كانت البطاقات ليس بها أي ثقب (بسبب نسيانها عند الثقيب) فتسقط في جيب خاص بذلك .

بعد انتهاء فرز البطاقات تبعاً للعمود الخاص بالجنس تعطينا الآلة حسب عدادها الخاص عدد البطاقات ذات الرقم ١ وهم الذكور وذات الرقم ٢ وهم الاناث . يمكننا بعد ذلك أن نعيد وضع البطاقات في الآلة لفرز البيان الخاص بالحالة الزوجية ، فاذا كان هذا البيان مثقب في العمود رقم ٦ نضبط المؤشر أمام هذا الرقم وندير الآلة فتدفع البطاقات في مجراها الخاص وتسقط بعضها في الجيب رقم ١ وبعضها في الرقم ٢ وبعضها في الرقم ٣ وبعضها في الرقم ٤ وبعضها في الجيب الأخير . وبذلك يعطينا عداد الآلة عدد البطاقات ذات الرقم ١ أي السكان العزاب ، والبطاقات رقم ٢ أي المتزوجين ، والبطاقات رقم ٣ أي المطلقين ، والبطاقات رقم ٤ أي الأراامل والبطاقات الأخيرة هم السكان دون سن الزواج ، وذلك اذا كانت هذه الأرقام هي الدليل الرقمي للحالة الزوجية .



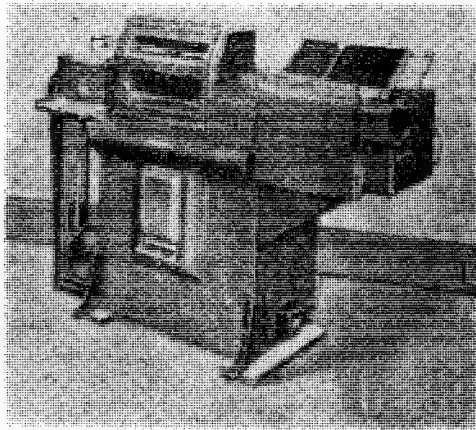
آلة الفرز الكهربائية

التبويب :

إذا كانت آلة الفرز تقوم بتصنيف البطاقات وعدها في كل فئة نريدها فما هي وظيفة آلة التبويب ؟ نفترض اننا فرزنا البطاقات الخاصة بالمستغلين وبذلك عرفنا عدد الذكور منهم وعدد الاناث ، ونريد بعد ذلك أن نعرف جملة الاجور التي تقاضاها كل من الجنسين ، نمرر البطاقات بعد فرزها داخل آلة كهربائية تسمى آلة التبويب . وهي تقوم بقراءة الأرقام المثقبة في الأعمدة الخاصة بالاجور في البطاقات ثم تحولها الى أرقام ثم تجمع هذه الأرقام مع بعضها بطاقة بعد أخرى وتسجل المجموع على قطعة ورق مثبتة في الآلة ويمكن نزعها منها . بهذه الطريقة نستطيع أن نعرف جملة الاجور التي تقاضاها كل من الجنسين للمستغلين مطبوعة ومبوبة على ورقة خاصة .

وآلة التبويب تستخدم لتحقيق اغراض مختلفة ، واستخدامها مبني على فكرة انها تستطيع أن تقرأ الثقوب الموجودة في أي عمود في البطاقة المثقبة ثم تحولها إلى الأرقام الخاصة بها ثم تطبعها على ورق خاص ، فاذا كنا نريد أن نحصل على بيان عن العمر والاجر لكل مشغل تبعاً للجنس ، نقوم بتصنيف البطاقات تبعاً للجنس في آلة الفرز ثم نمررها في آلة التبويب فنقوم

(طبقاً لنظام خاص) بتسجيل البيان الخاص بالعمر والاجر امام كل عامل في كل جنس من الجنسين ، أي انها تطبع لكل جنس عمودين احدهما يحتوي على العمر والآخر يحتوي على الأجر وبذلك تحصل على ورقة عليها عمر واجر كل عامل من الجنسين . ولذلك يمكن أن تستخدم هذه الآلات في اعداد القوائم الخاصة بمرتبات الموظفين في أي مؤسسة يعمل بها عدد كبير جداً من المشتغلين . وتستطيع آلة التنبؤ معالجة ٣٥٠٠ بطاقة في الساعة كما انها تستطيع تسجيل سبعة مجاميع مختلفة في نفس الوقت ، ولذلك فهي تسهل علينا العمل كثيراً وتساعدنا في الحصول على النتائج بسرعة فائقة .



آلة التنبؤ الكهربائية

عرض البيانات في جداول :

بعد جميع المعلومات وتصنيفها سواء باليد أو استخدام الآلات الإحصائية تكون الخطوة التالية في سلسلة العمليات الإحصائية هي عرض هذه المعلومات . وهناك طريقتان للعرض يمكن استعمالهما هما الجداول الإحصائية والرسوم البيانية .

والجدول الإحصائي يختلف عن أي كشف آخر، فالكشف يظهر معلومات

فردية أي عن كل وحدة على حدة بينما يظهر الجدول الاحصائي المعلومات بعد تجميعها وتصنيفها في فئات معينة قد تكون فئات جغرافية أو زمنية أو وصفية أو كمية ، وقد تكون تهجيناً لهذه الأنواع بحيث تظهر المعلومات في الجدول مصنفة تبعاً لنوعين من هذه الفئات أو أكثر . وبذلك عند عرض البيانات في الجداول الاحصائية تضيع معالم الشخصية وتصبح مجرد رقم معين في فئة معينة .

وبذلك يكون الجدول الاحصائي هو حلقة الاتصال بين الهيئة الاحصائية التي أعدته وبين كل من سوف يطلع عليه في المستقبل لأخذ المعلومات التي يريدها ، لذلك يجب أن يفسر نفسه بنفسه فلا يحتاج بذلك الى أي استفسار . لهذا لا بد أن تتوفر في الجدول الاحصائي شروط معينة حتى يكون واضحاً لا يشوبه أي إبهام أو أي شك في محتوياته .

عناصر الجدول الاحصائي وشروطه :

عند تصميم أي جدول احصائي يجب مراعاة القواعد الآتية :

أولاً - العنوان :

يجب أن يكون للجدول عنوان يجيب على الأسئلة الآتية :

أ - ما هي الأرقام التي يحتوي عليها الجدول ؟

ب - أي مكان أو أماكن تتعلق بها هذه الأرقام ؟

ج - أي فترة زمنية تتعلق بها هذه الأرقام ؟

د - ما هو الأساس أو الأسس التي صنفت تبعاً لها الأرقام في هذا الجدول ؟

مثال

الخيارات الزراعية في لبنان موزعة تبعاً لموقعها عام ١٩٦٥ .

وعند اعداد عنوان الجدول يجب مراعاة ان يظهر في اعلا الجدول ،
واذا كان طويلا يحسن كتابته في شكل هرم مقلوب ، كما يجب ان يكون
دقيقاً فيما يختص بالارقام الموجودة فيه :

مثال

الوفيات المسجلة في مكاتب وزارة الصحة في سوريا موزعة
تبعاً لمناطق الدولة وسبب الوفاة عام ١٩٦٥

ونلاحظ اننا كتبنا العنوان لأنه طويل في شكل هرم مقلوب ، واضفنا
الفقرة - المسجلة في مكاتب وزارة الصحة - حتى لا تؤخذ الارقام على انها تمثل
جميع الوفيات في الدولة . واذا كانت الارقام التي يظهرها الجدول هي ارقام
قياسية يجب أن يظهر تحت العنوان فترة الاساس لهذه الارقام :

مثال

الارقام القياسية لنفقة المعيشة في
الجمهورية العربية المتحدة
(عام ١٩٣٥ = ١٠٠)

ثانياً - عناوين الاعمدة :

كل عمود في الجدول يجب ان يكون له عنوان يوضح الارقام والبيانات
المدونة فيه توضيحاً كاملاً وبإيجاز . واذا لاحظنا ان تركيز عنوان أي عمود
في كلمات قليلة يجعله غير واضح يجب توضيحه توضيحاً كاملاً في ملاحظة
خاصة أسفل الجدول ، مثلاً قد يكون عنوان العمود عدد السكان ولكننا

نريد ان نوضح ان هذه الاعداد لا تشتمل على البدو والسكان الذين وجدوا خارج الدولة وخارج مياهاها الاقليمية وقت اجراء التعداد . ولكننا لا نستطيع أن نكتب كل هذا التوضيح في عنوان العمود ، لذلك نضطر الى تدوينه في اسفل الجدول .

ثالثا : عناوين الاسطر في كعب الجدول .

التوضيح الوصفي للارقام في كل سطر يكتب في العمود الاول من الجهة اليمنى من الجدول الذي يسمى بكعب الجدول . وعند تنظيم كعب الجدول يجب ملاحظة أن نكتب في قمة الكعب البنود الاكثر اهمية ، واذا كانت هذه البنود هي فترات زمنية يجب أن نبدأ بالفترات الاقدم ثم نتسلسل زمنياً حتى نصل الى أحدث فترة يظهرها الجدول . كذلك اذا كانت البيانات في كعب الجدول هي سنوات يحسن تقسيمها الى مجموعات خمسية يفصلها عن بعضها فراغات حتى تسهل قراءتها وحتى تبدو في مظهر أنيق . أما اذا كانت شهور يكون من الافضل تقسيمها الى مجموعات تتكون كل مجموعة منها من ثلاث شهور . واذا كان التوضيح في كعب الجدول يقسم الارقام الى مجموعات رئيسية واخرى فرعية تابعة لها يجب ترك بعض الفراغ تحت المجموعة الرئيسية ابتداء من اليمين قبل كتابة المجموعات الفرعية التابعة لها . كذلك اذا كان عنوان الارقام في سطر معين في الجدول يحتاج الى أكثر من سطر واحد تكتب الارقام أمام السطر الاخير في العنوان وليس امام الأسطر الأولى .

رابعا - المجاميع الرئيسية والفرعية :

يجب أن يظهر الجدول المجاميع الخاصة بالارقام الموجودة فيه كلما كان ذلك أمراً ممكناً سواء كان ذلك خاص بمجاميع أرقام الاعمدة أو ارقام

الأسطر . وإذا احتوى الجدول على مجاميع رئيسية وأخرى فرعية يجب كتابتها في الجدول بطريقة تساعد على الاهتمام اليها والتمييز بينها بسهولة . ويلاحظ أن تكتب المجاميع في نهاية الأرقام المكونة لها إلا إذا كان المطلوب أن يعطى الجدول الأهمية الأولى للمجاميع وبذلك يمكن كتابتها في قمة الأرقام المكونة لها . كذلك يجب التحقق من صحة المجاميع حيث يكون أمراً مخزياً أن يجد من يطلع على الجدول خطأ فيها إذ أن ذلك يجعله يفقد الثقة في جميع الأرقام الواردة في الجدول حيث أنه لن يستطيع أن يتبين مصدر الخطأ . وإذا كان هناك تقريب في الأرقام بحيث يكون هناك بعض الفرق بينها وبين مجاميعها يجب الإشارة إلى ذلك تحت الجدول في ملاحظة خاصة .

خامساً - المتوسطات والنسب المئوية :

إذا احتوى الجدول على المجاميع والمتوسطات يجب أن تسبق المجاميع المتوسطات . وإذا كانت الأرقام في عمود ما في الجدول عبارة عن نسب مئوية يجب أن نشير في عنوان العمود إلى نوع الرقم المنسوب إليه وبذلك يجب تجنب ذكر العلامة % فقط إذ يجب أن نذكر % من المجموع أو % من الزيادة أو % من النقص . . وهكذا . كذلك يجب أن يكون مجموع النسب ١٠٠ % إذا كانت النسب حسبت من المجموع ولكن قد يكون المجموع الفعلي للنسب يزيد أو ينقص ١٠٠ بسبب التقريب ، وفي هذه الحالة يكتب المجموع ١٠٠ % بحيث لا يزيد المجموع الفعلي للأرقام عن ١٠٠,٢ ولا ينقص عن ٩٩,٨ ، فإذا زاد الفرق عن ٠,٢ % يجب مراجعة العمليات الحسابية الخاصة بالنسب المئوية حتى نتأكد من عدم وجود أخطاء .

سادساً - الوحدات :

إذا كانت الأرقام في جميع الأعمدة مقاسة بنفس الوحدة يكتب نوع الوحدة بين قوسين تحت العنوان العام للجدول . أما إذا اختلفت وحدات

الارقام من عمود الى آخر أو من سطر الى آخر يكتب نوع الوحدة الخاصة بالعمود مع عنوانه والوحدة الخاصة بكل سطر مع عنوانه في كعب الجدول .
واذا اختلفت وحدات الارقام في عمود معين من سطر الى آخر يحسن أن يظهر الجدول عموداً خاصاً مجاوراً تظهر فيه وحدات الارقام كل على حدة (يظهر ذلك في جداول التجارة الخارجية الخاصة بكميات البضائع حيث تكون وحدة الكمية لكل نوع مختلفة عن الاخرى) .

سابعا - تنوير الارقام :

اذا كانت جميع الأرقام في الجدول مدورة أي محذوف منها نفس العدد من الارقام يكتب ذلك في العنوان العام للجدول ، فاذا كانت جميع الارقام مختصرة على مليون مثلاً يكتب ذلك في العنوان العام للجدول . واذا كانت الارقام في عمود معين فقط محذوف منها نفس العدد من الارقام يكتب ذلك في عنوان العمود نفسه (مثلاً الف دينار) . واذا كان الارقام في سطر معين محذوف منها نفس العدد من الارقام يكتب ذلك في التوضيح الخاص بالسطر في كعب الجدول .

ثامناً - الملاحظات :

تكتب الملاحظات التوضيحية للارقام أو العناوين في الجدول في أسفله . فاذا كانت هذه الملاحظات عامة بالنسبة لجميع الارقام في الجدول تكتب في أسفله دون أية اشارة . اما اذا كانت خاصة ببعض الارقام أو ببعض العناوين يجب وضع اشارة معينة على الرقم أو العنوان وكتابة نفس الاشارة أمام الملاحظات الخاصة بذلك اسفل الجدول ، واذا تعددت الملاحظات يجب استعمال اشارات مختلفة ، وغالباً تستعمل الارقام كإشارات مع العناوين والأشكال الهندسية كإشارات مع الارقام . وبشكل عام لا يجب أن يترك أي عنوان

أو أي رقم غامضاً أو به بعض الشك إذ يجب دائماً التوضيح والتفسير في ملاحظات أسفل الجدول .

وهناك اشارات اتفق عليها من قبل الهيئات الدولية يمكن ان تستعمل للدلالة على بعض المفاهيم مثل :

- وتدل على ارقام غير متوفرة .
- _____ وتدل على ارقام غير موجودة أصلاً
- × وتدل على ارقام غير رسمية .
- [] وتدل على ارقام غير محسوبة ضمن المجموع .

وبالرغم من أن هناك اتفاق عام على معنى هذه الاشارات ، الا انه يجب عدم اغفال ذكر دلالتها في مقدمة النشرة .

تاسعاً : المصدر

يجب الإشارة إلى مصدر الارقام في اسفل الجدول تحت الملاحظات اذا لم يكن مصمم الجدول المسئول عنها اصلاً . وعند ذكر المصدر يجب كتابه اسم المؤلف او الهيئة المسئولة اولاً ثم عنوان الكتاب أو النشرة أو المقال أو المجلة ثم اسم الناشر ثم تاريخ النشر ثم رقم الصحيفة الموجودة بها الارقام ويستحسن وضع عنوان المصدر بين قوسين من اعلا .

عاشراً : الخطوط الافقية التي يحتوي عليها الجدول .

اذا كانت الجاميع واردة في اول الجدول او ضمن جسمه يحسن فصلها عن الارقام التي تأتي بعدها بخطين متجاورين ، ولا يجب أن تقطع هذه الخطوط كعب الجدول ، وفيما عدا ذلك لا يجب أن يظهر في الجدول من الخطوط الافقية غير الخطوط الخاصة بعناوين الاعمدة وبالجاميع في نهاية الجدول من أسفل .

ملاحظات عامة :

١ - يحسن ترقيم الجداول التي تحتوي عليها النشرة الاحصائية حتى يمكن الاهتمام اليها بعد الاطلاع على فهرس الجداول في مستهل النشرة .

٢ - اذا كانت جميع الجداول في النشرة خاصة بنفس الدولة وب نفس الفترة الزمنية لا يجب أن نذكر ذلك في عنوان كل جدول اذ يكفي كتابتها مع العنوان العام للنشرة . ويجب أن لا ننسى اسم منطقة معينة أو فترة زمنية معينة في عنوان جدول معين اذا لم تكن الأرقام في هذا الجدول خاصة بالدولة عامة أو بالفترة الزمنية العامة للنشرة .

٣ - يمكن ترقيم عناوين الأعمدة وعناوين الأسطر في جدول معين حتى يسهل شرح محتوياتها ويكون ذلك أمراً واجباً في التقارير الاحصائية التي تحتوي على جداول وشرح للأرقام الواردة فيها .

٤ - يجب أن يكون الشكل العام للجدول مقبولاً فلا يجب أن يكون طويلاً وضيقاً أو قصيراً جداً وواسعاً كما لا يجب أن تظهر الأرقام فيه متضاربة ومعقدة النظر ، الأمر الذي ينفر القاريء . ويجب أن تظهر الأرقام في كل عمود مرتبة تحت بعضها من ناحية الخانات المختلفة التي تتكون منها .

التوزيع التكراري :

من أهم الخدمات التي يقدمها الاحصاء للابحاث المختلفة كيفية تنظيم واختصار البيانات بشكل يسمح للعقل أن يفهمها ويقف على أهم خصائصها . ومن أهم الوسائل التي يستخدمها الاحصائيون لهذا الغرض هو عمل توزيع تكراري لتلك البيانات . فلو فرض أن كان لدينا قيماً كثيرة لأطوال مجموعة من الأفراد وليكن عددهم ألفان مثلاً فليس من السهل على العقل أن يستوعب مميزات طول هؤلاء دون تلخيص وتنظيم مبدئي .

فيصعب على القامىء أن يعرف أصغر القيم وأكبرها لأطوال هؤلاء دون ضياع بعض الوقت والجهد ، وبالمثل يصعب معرفة ما اذا كان هناك تركيز في الأطوال عند قيم معينة بمجرد النظر الى القيم الدالة على أطوال هؤلاء الاشخاص ، وكذلك يصعب بمجرد النظر الى القيم الاصلية معرفة ما اذا كان هناك استمرار في التغير أي معرفة ما اذا كان هناك أشخاص يمثلون جميع الأطوال بين حدي المدى .

اذا فرض ان كانت البيانات الاصلية بمحدودة العدد - حوالي العشرة أو العشرين قيمة مثلاً - فمن الجائز أن نضعها في شكل مرتب اما في وضع تصاعدي أو تنازلي . ويطلق على القيم في وضعها الجديد اسم التوزيع المنتظم . فاذا أعطينا القيم الآتية وهي تدل على اجور عشرة من العمال في اليوم : ١٥ ، ١٠ ، ٤٠ ، ١٩ ، ١٨ ، ٣٠ ، ١٦ ، ٢٥ ، ٢٠ ، ٢٠ ، فقد يحسن ترتيب تلك الأجور ترتيباً تصاعدياً لكي يقف القارىء أولاً على مدى تغير القيم المعطاة (Range) وثانياً تركيزها عند قيمة أو قيم معينة أو انعدام هذا التركيز (Concentration) وثالثاً على استمرار هذه القيم خلال المدى كله (Continuity) أو انعدام هذا الاستمرار في بعض أجزائه وستكون القيم بعد ترتيبها هكذا .

١٠ ، ١٥ ، ١٦ ، ١٨ ، ١٩ ، ٢٠ ، ٢٠ ، ٢٥ ، ٣٠ ، ٤٠

فواضح ان مدى التغير في الاجور هؤلاء العمال يتراوح بين ١٠ - ٤٠ ليرة في اليوم وأن عدداً لا بأس به من أفراد المجموعة يتركز أجره عند ١٥ ليرة الى ٢٠ ليرة وأنه ليس هناك قيم بين ٣٠ - ٤٠ ليرة وهي نهاية المدى . ومن الواضح انه لم يكن من السهل الوصول الى تلك الحقائق دون استخدام فكرة التوزيع المنتظم . ولكن لا يمكن الاكتفاء بهذه الفكرة كوسيلة لاختصار البيانات إذا زاد عددها زيادة كبيرة فلا بد في مثل تلك الحالات من اللجوء الى التوزيع التكراري .

والفكرة الأساسية في التوزيع التكراري هي تحديد عدد مرات ظهور أي قيمة معينة ثم تسجيل تلك القيم في جدول وأمام كل منها عدد مرات ظهورها . ولتوضيح تلك الفكرة عملياً نفترض انه طلب الينا اختصار البيانات الآتية وهي عبارة عن أجور ٢٥٠ عامل في مصنع ما :

١٠ ١١ ١٤ ٢٠ ٢٥ ٣٥ ٢٩ ٣٦ ٤٥ ٢٤ ٢٠ ٢٣ ٣٤ ٤٣ ٢٨ ١٣ ١٢ ٨ ٢٠ ١٦
 ٢٧ ١٥ ٢٦ ٢١ ١٨ ٣١ ٣١ ٢٦ ٢١ ١٨ ٢٢ ٨ ١٥ ١٨ ٢٠ ٢٤ ٣٢ ٤٠ ١٢ ١٦
 ٣٥ ٣٣ ١٦ ٥٠ ٣٧ ٤ ٧ ١٠ ١٩ ٣٩ ٢٣ ٢٥ ٦٩ ١٩ ١٨ ١٥ ٢١ ١٧ ١٥ ٢٠
 ٢٧ ٢٤ ٢٢ ٢٠ ١٨ ١٦ ١٢ ٤٥ ٣٦ ٢٧ ٣٠ ١٧ ١٧ ١٥ ٣٨ ٤٥ ٦ ٩ ١٦ ٢٧
 ١٦ ١٩ ٢١ ٢٥ ٣٠ ٢١ ١٤ ١٩ ١٧ ١١ ١٥ ١٨ ٢٢ ٢٧ ١٥ ١٨ ٢٢ ٤٢ ٣٧ ٣٣

٣٦ ٣٠ ٣٨ ١٨ ٣٠ ١٣ ٢٥ ١٠ ٢١ ١٥ ١٥ ٢٣ ٣٥ ٣٠ ٢٥ ١٨ ١٤ ١٧ ١٩ ٢٠
 ٢٣ ١٣ ٢٠ ١٦ ١٠ ٣٨ ١٦ ٢٩ ١٨ ٢٥ ٣٠ ٣٥ ٢٨ ٢٧ ٢٥ ٢٢ ١٧ ٣٤ ٢٩ ٢٥
 ١٩ ١٥ ٢٠ ٢٦ ٣٠ ٢١ ٢٥ ١٩ ٢٠ ١١ ٣٨ ٣٢ ٢٧ ٥ ١٨ ٢٠ ١٦ ٢٠ ٣٢ ٣٦
 ٢٠ ١٧ ١٧ ١٥ ٢٥ ١٩ ١٠ ٢١ ٣٥ ١٧ ٢٥ ١٥ ٢٣ ٢١ ١٩ ١٧ ١٦ ١٣ ٢٥ ١٤
 ٣٥ ٣٥ ٢١ ١٥ ٥ ٨ ١٢ ٢٠ ٢٨ ٣٣ ٣٥ ٢٥ ١٩ ١٧ ٩ ٢٤ ٢٧ ٣٥ ٢٩ ١٩

١٥ ١٤ ١٧ ٣١ ١٩ ٢٠ ٢٦ ٢٥ ٣٥ ٣٠

٥ ٢١ ١٩ ٣٠ ٣٠ ١٢ ١٨ ٢٢ ٢٢ ٢٢

٢١ ٢٥ ٣٠ ١٠ ١٤ ١٦ ٢٣ ٢٠ ١٩ ١٩

٢٤ ٤١ ٤٠ ٤٠ ١٨ ٢٥ ٢٣ ١٩ ١٥ ١٧

١٩ ٢٥ ٣٠ ٢٨ ١٥ ٢٧ ٢٨ ٣٤ ٢٩ ١٥

وطريقة عمل التوزيع التكراري وفقاً لتعريفه هي تسجيل قيم المتغير واحدة بعد أخرى حسب ورودها في الكشف الأصلي ثم وضع علامات امام

كل قبعة يتكرر حدوثها وذلك لتحديد عدد مرات تكرارها ويسهل بعد ذلك ترتيب القيم ترتيباً تصاعدياً أو تنازلياً وأمام كل قيمة تكرارها .

والواقع ان هذه الطريقة المبدئية في عمل التوزيع التكراري لا تفيدنا كثيراً إذ ان تكرارات القيم المختلفة ستزايد وتتناقص بطبيعة الحال بلا نظام ولذا يجب ضم بعض القيم الى البعض الآخر وتكوين فئات (groups) أو (Classes) فلو أخذنا العمال الذين يتقاضون أجور تتراوح بين ٨،٤ ليرة لحصلنا على تكرار قدره ٦ لهذه الفئة ، وبالمثل لو أخذنا الذين يتقاضون بين ٨ ، ١٢ لحصلنا على تكرار قدره ١٤ لتلك الفئة . وهكذا اذا كررنا عملية ضم القيم وعد تكراراتها لكان الناتج جدولاً تركزياً أنسب في صورته لتحقيق الغرض الذي من أجله قمنا بعمل التوزيع التكراري .

فئات الاجور بالليرة التكرار (عدد العمال)

٤ -	٦
٨ -	١٤
١٢ -	٣١
١٦ -	٥٤
٢٠ -	٤٣
٢٤ -	٣٦
٢٨ -	٢٧
٣٢ -	١٩
٣٦ -	١٠
٤٠ -	٦
٤٤ -	٣
٤٨ - ٥٢	١

ويستطيع القارئ أن يحس بنفسه فائدة التوزيع التكراري في اختصار البيانات فقد كان يستحيل معرفة مدى تغير الأجور دون بذل مجهود ووقت في حين أن ذلك واضح لأول وهلة من مجرد قراءة التوزيع التكراري .

وبالمثل لا يتضح تركيز القيم عند أي قيمة معينة أو انعدام هذا التركيز لمن يكتفي في دراسته بالكشف الأصلي ولكن على العكس من ذلك واضح جلي أن الأجور تتركز عند الأجر من ١٦ الى ٢٠ ليرة .

وكذلك لا يمكن الجزم باستمرار التغير في الأجور في المدى كله أو انعدام هذا الاستمرار لمن يطلع على كشف الأجور في حين أنه من أسهل الأمور الحكم على هذا بالنظر في جدول التوزيع التكراري حيث يتضح استمرار التغير بين ٤٠ ليرة ، ٥٠ ليرة . تلك الحقائق الثلاث ، تحديد المدى والتركز أو انعدامه والحكم على استمرار التغير خلال المدى أو عدمه من أهم الفوائد العملية للتوزيعات التكرارية .

وليس من الضروري عند عمل التوزيع التكراري أن يحدد عدد مرات تكرار كل قيمة على حدة ثم ضم القيم في فئات كما ذكرنا إذ كان الغرض من وصف تلك الطريقة تأكيد الفكرة الأساسية التي يبنى عليها أي توزيع . أما الطريقة العملية لذلك فهي تحديد مدى التغير للظاهرة المراد وصفها ثم تقسيم هذا المدى الى فئات يستحسن أن تكون متساوية في الطول ووضعها في جدول تقريبي كما هو موضح في الجدول الآتي .

جدول تفرغ أجور ٢٥٠ عامل .

فئات الأجور	التفرغ	التكرار (عدد العمال)
٤ -	1 IIII	٦
٨ -	IIII IIII IIII	١٤
١٢ -	IIII IIII IIII IIII IIII	٣١
١٦ -	IIII IIII IIII IIII IIII IIII IIII	٥٤
٢٠ -	IIII IIII IIII IIII IIII IIII IIII IIII	٤٣
٢٤ -	IIII IIII IIII IIII IIII IIII IIII IIII	٣٦
٢٨ -	IIII IIII IIII IIII IIII IIII IIII IIII	٢٧
٣٢ -	IIII IIII IIII IIII IIII IIII IIII IIII	١٩
٣٦ -	IIII IIII IIII IIII IIII IIII IIII IIII	١٠
٤٠ -	1 IIII	٦
٤٤ -	IIII	٣
٤٨ - ٥٢	1	١
المجموع		٢٥٠

والخطوة التالية لذلك قراءة الأرقام الاصلية قيمة ووضعية علامة (//) أمام كل فئة تقع بين حدودها تلك القيمة واحدة بعد الأخرى .

وتسهيلاً لعملية العد النهائي يستحسن وضع العلامات في مجموعات خماسية (IIII) كما هو واضح ولا يتبقى بعد ذلك الا حصر تلك العلامات وتسجيل عددها أمام كل فئة . ويسمى الجدول الذي استخدم في توزيع القيم في فئاتها جدول تفرغ (Tallying Table) ، أما الجدول الذي يوضح الفئات والتكرارات دون العلامات المساعدة يسمى جدول تكراري للاختصار . ويمكن تطبيق فكرة التوزيعات التكرارية على البيانات غير الرقمية اسوة بالرقمية منها . فواضح أن الفكرة الأساسية في أي توزيع تكراري هي تحديد عدد مرات ظهور القيم المختلفة للظاهرة موضوع الدراسة .

فلو فرض أن كان لدينا كشفاً بالحالة الزوجية للعمال السابقين فيمكن تلخيص تلك البيانات في شكل توزيع تكراري غير رقمي بكتابة الحالات

الزواجية المختلفة ثم تحديد عدد العمال الذين يشتركون في كل حالة معينة كما فعلنا في مسألة الأجور بالضبط. ويكون التوزيع التكراري المطلوب هكذا:

الحالة الزوجية	التكرار (عدد العمال)
أعزب	٤٨
متزوج	١٣١
مطلق	٤١
أرمل	٣٠

ويمكن وضع التوزيع التكراري المطلق في شكل توزيع نسبي وذلك بحساب نسبة تكرار كل فئة من مجموع التكرارات .

التوزيع التكراري المستمر وغير المستمر - المتصل وغير المتصل :

ويجدر بنا أن نشير الى وجود نوعين أساسيين من المتغيرات (Variables) فهناك متغيرات يمكن بطبيعتها أن تأخذ جميع القيم بين حدي مدى التغير ويسمى هذا النوع « متغيرات مستمرة » (Continuous Variables) ، ومثله الأجور في المثال الذي أشرنا اليه إذ ان الأجر يمكن - نظرياً على الأقل - أن يأخذ جميع القيم بين ٤ ليرات ، ٥٢ ليرة وهو مدى التغير في المثال السابق - فيمكن أن نتصور وجود عامل أجره $\frac{1}{7}$ ٤ ليرة وثالث أجره $\frac{3}{4}$ ٤ ليرة ورابع أجره ٥ ليرة وخامس أجره $\frac{1}{3}$ ٥ ليرة وهكذا حتى نهاية المدى .

وبالمثل اذا كان لدينا توزيعاً تكرارياً لأوزان ألف طالب فان المتغير وهو وزن الطالب متغير مستمر حيث أن وزن كل طالب منهم يمكن ولو نظرياً على الأقل أن يأخذ أي قيمة بين حدي مدى التغير .

وكذلك اذا كان لدينا توزيعاً تكرارياً للأسعار الشهرية للقطن خلال العشرين سنة الماضية فان المتغير هنا وهو السعر متغير مستمر إذ لو تبين لنا

أن السعر تغير في المدى بين عشرين ريال ، ٢٠٠ ريال مثلا يمكن تصور حدوث أي سعر بين هذين الحدين خلال المدة موضوع الدراسة .

وعلى العكس من ذلك يوجد نوع آخر من المتغيرات لا يمكن بطبيعته أن يأخذ جميع القيم التي تقع في حدود مدى التغير وهذا النوع يُطلق عليه اسم المتغيرات غير المستمرة أو الوثابة (Discontinuous Variables) .

فلو فرض ان كان لدينا توزيعا تكراريا لعدد أفراد أسر حي معين من أحياء بيروت ، فالمتغير هنا هو عدد أفراد الأسر متغير وثاب ، حيث أن عدد أفراد الأسرة لا يمكن أن يأخذ جميع القيم بين حدي المدى . فلو فرضنا أن عدد الأفراد أسر هذا الحي يتراوح بين ١ ، ٨ أفراد مثلا ، فيمكن تصور وجود عائلات عدد أفرادها ٢ ، ٣ ، ٤ ، وهكذا . ولكن لا يمكن تصور وجود عائلات عدد أفرادها $1\frac{1}{2}$ أو $1\frac{1}{2}$.

ولو كان لدينا توزيعا تكراريا لعدد ثمار التفاح الذي تحمله اشجار صنف معين منها وثبت لنا أن مدى التغير يتراوح بين ٥٠ ثمرة ، ٢٠٠ ثمرة مثلا فالمتغير هنا هو عدد الثمار متغير وثاب اذا لا يمكن تصور وجود شجرة تحمل $1\frac{1}{2}$ ٧٠ ثمرة أو $1\frac{1}{4}$ ٩٣ ثمرة مثلا

ومن البديهي أن جميع التوزيعات التكرارية للقيم غير الرقمية لا يمكن أن يكون المتغير فيها مستمرا ، ففي مثال تقسيم العمال حسب الحالة الزوجية كان المتغير وثابا بطبيعته ولا يمكن أن يكون غير ذلك .

ونلاحظ ان توزيع الطلبة تبعاً للاعمار وان كان يمكن ان يكون توزيعا متصلا ، الا انه حيث أن الاعمار تكون بالسنوات الكاملة لذلك تكتب فئات العمر على أساس غير متصل .

التوزيع التكراري المفتوح والمغلق :

والتوزيعات التكرارية نوعان الأول مفتوح وهو ما لم يتعين حده الاعلى

أو حده الأدنى أو كليهما لفئتيه المتطرفتين. والثاني مغلقي وهو ما عرف حدا فئتيه العليا والسفلى ومثل ذلك التوزيع السابق لأحور العمال حيث أن الحد الأدنى للفئة الاولى والحد الأعلى للفئة الأخيرة معروفان وهناك حالات تتطلب طبيعة البيانات فيها عدم تعيين الحد الأدنى للفئة الأولى منها أو الحد الأعلى للفئة الأخيرة اما لاعتبارات عملية حيث تقتضي الإشارة إلى كثير من المفردات غير المتجانسة بأنها تقع في فئة غير محدودة الأطراف ، واما لجهل الباحث بمحدود الفئات فمثلا قد لا يستطيع ممثلوا وزارة الزراعة معرفة المساحات التفصيلية التي يملكها صغار الملاك في قرية معينة على الرغم من معرفة عددهم ، ولذا يطلقون على الفئة الاولى أقل من نصف فدان مثلا جهلا بحقيقة الحد الأدنى للفئة .

ومن الأفضل أن يحاول الباحث جعل التوزيع التكراري الذي يقوم بعمله مقبول الطرفين كلما أمكن ذلك ، اذ ان جعله مفتوحا يترتب عليه عدم امكانية استخدام بعض المتوسطات - كما سنشير إلى ذلك عند الكلام عن المتوسطات - وكذلك يصعب وضع التوزيع التكراري المفتوح في شكل بياني كامل كما سيظهر عند معالجة هذا الموضوع فيما بعد .

التوزيع المنتظم وغير المنتظم :

إذا كان مدى الفئات في التوزيع التكراري متساوي في جميع الفئات فإن التوزيع يكون منتظماً . ولكن يبدو أحياناً ان القيم تتركز في جزء من التوزيع بينما تكون مبعثرة في الاجزاء الأخرى ، ولذلك إذا جعلنا التوزيع منتظماً فان كثير من الفئات تكون خالية من التكرارات ولهذا يستحسن أن يكون التوزيع غير منتظم يجعل الفئات غير متساوية وذلك بان تكون الفئات في الاجزاء التي تتركز فيها القيم ذات مدى قصير بينما تكون ذات مدى واسع في الاجزاء التي تتبعثر فيها القيم .

تحديد عدد الفئات :

من أصعب الأمور اختيار عدد الفئات التي يجب توزيع القيم الأصلية

عليها اذ ان ذلك يتوقف على اعتبارات عدة . هذا وان كانت الخبرة والمران خير مرشدين لأنسب الوسائل الا انه لا مانع من ذكر بعض الاعتبارات الأساسية التي لا يجب ان يغفلها الباحث في هذا الشأن .

فلا يجب ان نغالي في الأكتثار من عدد الفئات اذ ان ذلك يتنافى والسبب الرئيسي الذي من اجله قمنا بعمل التوزيع التكراري نفسه الا وهو الاختصار والتركيز . ومن الناحية الأخرى لا يجب ان نغالي في الاختصار نفسه ونضع التوزيع في فئتين أو ثلاث اذ ان في ذلك اختصار محل .

هذا وان كان من الصعب اعطاء قاعدة لهذا الموضوع الا انه لا يجوز على العموم ان يقل عدد فئات التوزيع عن عشرة ولا يزيد عن عشرين فئة تبعا لكثرة عدد القيم المراد تلخيصا أو قلتها .

ومما لا بأس به الإشارة هنا إلى قاعدة رياضية لتحديد عدد فئات التوزيعات التكرارية ذات القيم متوسطة العدد أي بين مائة وألف قيمة ويطلق على هذه القاعدة ، قاعدة ستيرجس Sturges Rule وهي :

عدد فئات التوزيع التكراري $= 1 + 3.3 \times \text{لوغارتم عدد القيم}$.
وعند تقسيم القيم الى فئات يجب ان نلتفت إلى قاعدتين أساسيتين ، الأولى هي الا تكون الفئات متباعدة بمعنى ان يكون هناك هوة أو مسافة بين الحد الأعلى للفئة والحد الأدنى للفئة التالية لها . فلو كنا في صدد تقسيم عدد من العمال حسب ساعات العمل التي يشتغلها كل واحد منهم اسبوعيا فلا يجب وفقا لتلك القاعدة - ان يكون التقسيم كالآتي :

٤٠ ساعة إلى ٤٥ ساعة

٤٦ ساعة إلى ٥٠ ساعة

٥١ ساعة إلى ٥٥ ساعة

إذا أن هناك مسافة بين الحد الأعلى للفئة الأولى والحد الأدنى للفئة الثانية وهكذا ، فيقال أن هناك تباعد بين تلك الفئات - فاذا اتبعنا هذه الطريقة

في تقسيم الفئات يكون من الصعب تحديد الفئة التي ينتمي اليها عامل عدد ساعات عمله الاسبوعية ٤ ساعة ونصف أو آخر عدد ساعات عمله ٥٠ ساعة وربع . هذا ولا يجب أن نفعل عند تطبيق هذه القاعدة طبيعة المتغير نفسه ودرجة التقريب فيه . فلو كان المتغير وثاباً مثلاً فما لا غبار عليه أن نجعل الفئات متباعدة إذ انه من الخطأ أن نجعل الفئات غير ذلك .

فلو كنا بصدد تقسيم عدد كبير من الشقق بحسب عدد الغرف التي تحتوي عليها كل منها فلا مانع بل لا بد من جعل الفئات متباعدة هكذا :

غرفتان إلى أربع غرف
خمسة غرف إلى سبعة غرف
ثمانية غرف إلى عشرة غرف

فقد يظن القارئ لأول وهلة أن هناك تباعد بين تلك الفئات خلافاً للقاعدة ولكنه لو تذكر أن المتغير هنا وثاب وأنه لا يمكن أن يأخذ قيما تقع بين أي من تلك الفئات المذكورة لتحقيق من صحة التقسيم .

وبالمثل قد يكون السبب في التباعد بين الفئات التقريب الحسابي في قيم المتغير الأصلي فلو فرض أننا نقوم بعمل توزيع تكراري لطول محيط صدر مجموعة من الطلبة بالسنتيمتر مثلاً وأن محيط الصدر كان يقاس بأقرب سنتيمتر فلا مانع اذا من وضع الفئات هكذا :

١٢١ سنتيمتر إلى ١٣٠ سنتيمتر
١٣١ سنتيمتر إلى ١٤٠ سنتيمتر
١٤١ سنتيمتر إلى ١٥٠ سنتيمتر

فلا يمكن قبول الانتقاد الذي سبق أن وجهناه للفئات المتباعدة في هذه الحالة إذ أن من كان محيط صدره ١٣٠ سنتيمتر يقع في الفئة الأولى ومن كان محيط صدره ١٣٠ سنتيمتر وكسر لا بد أن يسجل طول محيط صدره لأقرب

سنتمتر فاما ان يكتب ١٣٠ سنتمتر ويقع لذلك في الفئة الأولى أو ١٣١ سنتمتر ويقع لذلك في الفئة الثانية .

أما إذا كان المتغير مستمراً فيكون من الخطأ ان نجعل هناك أي تباعد بين الفئات بمعنى ان الحد الأدنى لكل فئة هو نفسه الحد الأعلى للفئة السابقة .

والقاعدة الثانية التي يجب أن نتبعها عند تقسيم القيم إلى فئات هي ألا تكون الفئات متداخلة Overlapping Groups بمعنى أن يكون جزء من فئة داخل في فئة أخرى سابقة أو لاحقة . فلو كنا بصدد تقسيم الشركات المساهمة حسب معدلات للربح التي توزعها كل شركة على حاملي أسهمها وقسمنا التوزيع إلى الفئات الآتية :

٢٪ إلى ٥٪

٥٪ » ٨٪

٨٪ » ١١٪

فيختار الباحث في وضع الشركة التي وزعت ربحاً على مساهميها بنسبة ٥٪ بين الفئتين الأولى والثانية . وبالمثل من الخطأ اعتبار الشركة التي وزعت ربحاً بمعدل ٨٪ من بين وحدات الفئة الثانية دون الثالثة اذ ان هاتين الفئتين تتنازعاها وذلك بسبب تداخل الفئات بعضها في البعض الآخر .

طرق كتابة الفئات :

نظراً للاعتبارات السابقة يجب كتابة الفئات باحدى الطرق الآتية :

٢٪ وأقل من ٥٪

٥٪ » » ٨٪

٨٪ » » ١١٪ وهكذا

واختصاراً للكتابة جرت عادة الاحصائيين على كتابة هذه الفئات نفسها هكذا :

$$\begin{array}{rcl}
 \frac{2}{5} & - & \text{لقرأ } \frac{1}{5} \text{ وأقل منه } \frac{1}{5} \\
 \frac{2}{5} & - & \text{لقرأ } \frac{1}{5} \text{ وأقل منه } \frac{1}{5} \\
 \frac{2}{5} & - & \text{لقرأ } \frac{1}{5} \text{ وأقل منه } \frac{1}{5} \\
 \frac{2}{5} & - & \text{لقرأ } \frac{1}{5} \text{ وأقل منه } \frac{1}{5} \\
 \frac{2}{5} & - & \text{لقرأ } \frac{1}{5} \text{ وأقل منه } \frac{1}{5} \\
 \frac{2}{5} & - & \text{لقرأ } \frac{1}{5} \text{ وأقل منه } \frac{1}{5} \\
 \frac{2}{5} & - & \text{لقرأ } \frac{1}{5} \text{ وأقل منه } \frac{1}{5} \\
 \frac{2}{5} & - & \text{لقرأ } \frac{1}{5} \text{ وأقل منه } \frac{1}{5}
 \end{array}$$

ويمكن كتابة الفئات بصورة أخرى هكذا :

$$\begin{array}{rcl}
 \frac{2}{5} & \text{أكثر من } \frac{2}{5} & \text{إلى } \frac{5}{5} \text{ وتكتب باختصار } \frac{2}{5} \\
 \frac{2}{5} & \text{أكثر من } \frac{5}{5} & \text{إلى } \frac{8}{5} \text{ وتكتب باختصار } \frac{2}{5} \\
 \frac{2}{5} & \text{أكثر من } \frac{8}{5} & \text{إلى } \frac{11}{5} \text{ وتكتب باختصار } \frac{2}{5}
 \end{array}$$

هذا ولو ان تلك الطريقة صحيحة إلا أنها ليست شائعة الاستعمال كالطريقة الأولى .

في كل من الطريقتين مدى الفئة ٣ ومراكز الفئات ٣، ٥، ٦، ٩ . والفرق بينها ان الحد الأعلى للفئة الأولى مثلاً هو أقل بقليل من $\frac{5}{5}$ في الطريقة الأولى و $\frac{5}{5}$ تماماً في الطريقة الثانية - كذلك الحد الأدنى للفئة الأولى هو $\frac{2}{5}$ أي ان الرقم $\frac{2}{5}$ داخل في الفئة الأولى تبعاً للطريقة الأولى بينما في الطريقة الثانية الحد الأدنى للفئة يزيد قليلاً عن $\frac{2}{5}$ والصورة الثالثة لكتابة الفئات هي هكذا :

$$3,5$$

$$6,5$$

$$9,5$$

وهي تعني الفئات التي مراكزها ٣، ٥، ٦، ٩ . ويمكن معرفة المدى بحساب الفرق بين كل مركز وآخر وهو ٣ في هذه الحالة ، فاذا أردنا

$$\text{أن نحدد الفئات التي لها هذه المراكز نقسم المدى على } 2 \text{ أي } \frac{3}{2} = 1,5$$

بعد ذلك تطرح ١,٥ من كل مركز نحصل على الحد الأدنى للفئة ونجمع ١,٥

على كل مركز نحصل على الحد الأعلى للفئة وتبعاً لذلك تكون الفئات التي مراكزها الأرقام السابقة هي ٢ - ٥٠ - ٨٠ - ١١ .

ومن أهم ما يجب التنويه به ونحن في صدد اختيار فئات التوزيع التكراري العمل على جعل جميع الفئات متساوية المدى ما أمكن وسوف نرى عند الكلام على موضوع المتوسطات ان في تساوي جميع الفئات اختصار كبير في العمل الحسابي كما أن في عدم التساوي حرمان من امكان استخدام بعض أنواع المتوسطات الهامة .

ويجب أن نذكر بتلك المناسبة ما سبق أن أشرنا اليه وهو ان هناك حالات تستدعي طبيعة البيانات جعل الفئات غير متساوية . فمن أراد أن يدرس التوزيع التكراري للنسب المئوية للمتوفين في فئات العمر المختلفة بين سكان بلد معين بحسب العمر عند الوفاة لا يمكن أن يجعل التوزيع متساو في الفئات . فمن المعروف أن نسب المتوفين في الأعمار المتوسطة أي بين عشرين سنة وخمسين سنة تكاد تكون واحدة فلا يتطلب الأمر التفصيل الدقيق في تلك المرحلة من العمر ، ولكنه من المعروف أيضاً ان درجة تعرض الأطفال وخصوصاً الرضع منهم أي الذين لم يبلغوا الواحدة بعد سريعة التذبذب من سنة الى أخرى بل ومن شهر الى آخر ، الأمر الذي يستدعي من الباحث المدقق العمل على جعل فئات التوزيع قصيرة في أول التوزيع وطويلة في وسطه ثم صغيرة في نهايته . أما فيما عدا تلك الحالات وأمثالها يجب أن يحاول الباحث أن يختار الفئات بحيث تكون كلها متساوية . وهناك اعتبار آخر يجب الالتفات اليه عند اختيار فئات التوزيعات التكرارية وهو أن مركز كل فئة سيمثل جميع الأفراد الذين يقعون بين حدودها وهذا ما يتطلب منا يقظة وحذر شديد في اختيار تلك الحدود اذ لا بد من الالتفات الى طبيعة البيانات الأصلية وطريقة تركزها .

فمن أراد أن يكون توزيعاً تكرارياً لأعمار فئة كبيرة من السكان لا بد أن يتذكر تقريب الأرقام عند اعطاء السن ، فمن كان عمره الحقيقي ١٩ سنة

أو ٢١ سنة مثلاً سيجعله ٢٠ سنة . أما الذين أعمارهم ٣٧ سنة أو ٣٨ سنة أو ٣٩ سنة سيجعلونه ٤٠ سنة ولذا وجب اختيار الفئات كالآتي :

١٥	وأقل من	٢٥
٢٥	»	٣٥
٣٥	»	٤٥ وهكذا .

فواضح ان الذين اعتبروا في الفئة الأولى لأنهم سجلوا أعمارهم ٢٠ سنة هم في الواقع من ١٥ الى ٢٤ سنة فكان التقسيم الى تلك الفئات صحيح بقدر الأماكن الخطأ الغير المقصود في تسجيل العمر .

ومما هو جدير بالذكر أن مراكز الفئات يمكن تحديدها بحساب متوسط الحدين الأول والأعلى لكل فئة ، فمثلاً في التوزيع ٧ - ٩ - ١١ - ١٣ - ١٥ ،

يكون مركز الفئة الأولى $\frac{7+9}{2} = 8$ ومركز الفئة الثانية $\frac{9+11}{2} = 10$

وهكذا . وهذه القاعدة عامة نطبقها مهما اختلف نوع التوزيع النكراري مع ملاحظة ان التوزيع المفتوح لا يمكن حساب مركز فئته المفتوحة لعدم معرفة أحد حديها

التوزيعات التكرارية المتجمعة :

من التوزيع التكراري العادي يمكن معرفة عدد الوحدات التي تقع في داخل كل فئة مثلاً نعرف عدد الاشخاص الذين يتراوح طولهم بين ١٦٠ سم ، ١٦٥ سم . ولكن قد يطلب أحياناً معرفة عدد الطلبة الذين يقل طولهم عن ١٦٥ سم أو يزيد طولهم عن ١٦٠ سم ، ولا يمكن الاجابة بسهولة وبشكل مباشر على تلك الأسئلة بمجرد النظر الى التوزيع العادي . ولكي نتمكن من الاجابة على تلك الاسئلة وغيرها لا بد من وضع التوزيع في شكل جديد سنطلق عليه اسم التوزيع التكراري المتجمع .

والفكرة الأساسية في التوزيعات التكرارية المتجمعة هي تجميع التكرارات امام الحد الاعلى لكل فئة وفي هذه الحالة يكون التوزيع التكراري متجمع صاعد حيث أن التكرارات في صعود مستمر ، أو تجميعها أمام الحد الأدنى لكل فئة ابتداء من أسفل التوزيع وفي هذه الحالة يكون التوزيع متجمع هابط حيث تكون التكرارات في هبوط مستمر .

توزيعان تكراريان متجمعان صاعد وهابط

توزيع تكراري بسيط				توزيع تكراري متجمع	
فئات الأجر	التكرار (عدد العمال)	الحدود العليا للفئات	التكرارات المتجمع الصاعد	الحدود الدنيا للفئات	التكرارات المتجمع الهابط
٤ -	٦	أقل من ٨	٦	٤ وأكثر	٢٥٠
٨ -	١٤	٨ » ١٢	٢٠	٨ » »	٢٤٤
١٢ -	٣١	١٢ » »	٥١	١٢ » »	٢٣٠
١٦ -	٥٤	٢٠ » »	١٠٥	١٦ » »	١٩٩
٢٠ -	٤٣	٢٤ » »	١٤٨	٢٠ » »	١٤٥
٢٤ -	٣٦	٢٨ » »	١٨٤	٢٤ » »	١٠٢
٢٨ -	٢٧	٣٢ » »	٢١١	٢٨ » »	٦٦
٣٢ -	١٩	٣٦ » »	٢٣٠	٣٢ » »	٣٩
٣٦ -	١٦	٤٠ » »	٢٤٠	٣٦ » »	٢٠
٤٠ -	٦	٤٤ » »	٢٤٦	٤٠ » »	١٠
٤٤ -	٣	٤٨ » »	٢٤٩	٤٤ » »	٤
٤٨ - ٥٢	١	٥٢ » »	٢٥٠	٤٨ » »	١

والفائدة العملية لهذا النوع من التوزيع التكراري اننا نستطيع أن نعرف بمجرد النظر الى الجدول السابق عدد العمال الذين يتقاضون أجور أقل من أجر معين او أعلى من أجر معين ، فمثلا يتضح ان هناك ٥١ عامل يتقاضون اجراً اقل من ١٦ ليرة وان ١٩٩ عامل يتقاضون أجراً يزيد عن ١٦ ليرة وبالطبع يمكن معرفة نسبة هذا العدد الى المجموع ووضعها في خانة مقابلة للجدول .

كذلك من هذا الجدول يمكن معرفة الأجر الذي يتقاضاه عدد معين من العمال ، فمثلا من التوزيع المتجمع السابق يتضح ان الأجر الذي يتقاضاه ٢٣٠ من العمال هو أقل من ٣٦ ليرة ، كما يمكننا أن نعرف ان هذا العدد من العمال يتقاضى أكثر من ١٢ ليرة .

ونلاحظ انه إذا كان لدينا توزيع تكراري متجمع فيمكن بعملية عكسية (عملية طرح) أن نحصل على التوزيع التكراري الأصلي ، فمثلا :

الحد الأعلى للفئة	التكرار المتجمع الصاعد	الفئات	التكرار
أقل من ٦٤,٥	٣	- ٦٤,٥	٣
٧٤,٥ د د	٢٤	- ٦٤,٥	٢١
٨٤,٥ د د	١٩٢	- ٧٤,٥	٧٨
٩٤,٥ د د	٢٨٤	- ٨٤,٥	١٨٢
١٠٤,٥ د د	٥٨٩	- ٩٤,٥	٣٠٥
١١٤,٥ د د	٧٩٨	- ١٠٤,٥	٢٠٩
١٢٤,٥ د د	٨٧٩	- ١١٤,٥	٨١
١٣٤,٥ د د	٩٠٠	- ١٢٤,٥	٢١
١٤٤,٥ د د	٩٠٥	١٣٤,٥ - ١٤٤,٥	٥

كذلك نلاحظ أن التوزيع المتجمع المطلق سواء كان صاعداً أم هابطاً يمكن تحويله إلى توزيع متجمع نسبي وذلك بحساب نسبة كل تكرار متجمع من مجموع التكرارات الذي يكون هو آخر تكرار متجمع .

التوزيع التكراري المزدوج :

يذكر القارئ ان من أهم فوائد التوزيع التكراري تلخيص البيانات في شكل يسمح للعقل بالوقوف على أهم خصائص القيم التي نريد اختصارها . ولكن حتى الآن كنا نقتصر على معالجة عدد من قيم متغير واحد .

ولكن لو كان لدينا بيانات تخص ظاهرتين بينهما علاقة مثلاً بيانات عن عمر الزوج وعدد الأولاد الذين عنده ، أو عن الأسعار المختلفة لسلمة ما والكمية التي تعرض بهذه الأسعار ، أو عن علامات بعض الطلبة في اختبار الذكاء وعلاماتهم في اختبار علم من العلوم ، أو عن كمية المطر المتساقط في شهر يوليو ومتوسط كمية الانتاج من القمح ، أو عن أطوال مجموعة من الطلبة وأوزانهم - في مثل هذه الحالات لا يمكن اختصار البيانات بواسطة التوزيع التكراري العادي حيث يكون لدينا نوعان من البيانات أمام كل وحدة ، وقد تكون هذه البيانات مقاسة بوحدات مختلفة . فإذا حاولنا اختصار كل نوع منها في توزيع خاص بها سوف يكون لدينا توزيعان منفصلان عن بعضهما الأمر الذي لا يساعدنا في دراسة العلاقة بين الظاهرتين . ولهذا نحاول تبويبها في جدول واحد ، هو جدول التوزيع التكراري المزدوج ، فمثلاً إذا كان لدينا البيانات الآتية عن أجور بعض من العمال وعدد أولاد كل منهم فإنه من الممكن اختصار تلك البيانات في جدول توزيع تكراري مزدوج وذلك بتحديد الفئات (الأرقام) الخاصة بالعمر والفئات الخاصة بعدد الاطفال ثم نبدأ بتفريغ البيانات السابقة في الفئات المزدوجة التي تقابل بعضها ويكون التفريغ بوضع علامات في كل خانة تم عن الفئة المزدوجة التي ينحصر بين مداها كل زوج من القيم . بعد ذلك نلخص جدول التفريغ في جدول تكراري مزدوج بترجمة العلامات إلى أرقام كما فعلنا في حالة التوزيع التكراري البسيط تماماً .

2-20 '2-23 '0-32 '0-23 '0-29 '0-13 '0-12 '1- 8 '3-20 '0-17
 '0-10 '0-11 '0-12 '1-20 '0-20 '7-30 '1-29 '2-37 '3-20 '1-22
 '0-22 '0- 8 '1-10 '2- 18 '1-20 '0-22 '0-32 '2-20 '3-12 '0-17
 '2-27 '2-10 '3-27 '3-21 '7-18 '2-31 '2-31 '0-27 '1-21 '0-18
 '7-23 '2-20 '3-29 '7-19 '0-17 '2-10 '0-21 '0-17 '3-10 '2-20

 '1-10 '1-23 '2-30 '3-30 '3-20 '3-18 '0-12 '2 17 '0-19 '1-20
 '2-22 '2-20 '0-38 '2-18 '2-30 '1-13 '2-20 '0-10 '0-21 '0-10
 '3-30 '0-30 '0-28 '7-27 '0-20 '1-22 '1-17 '1-32 '3-29 '3-20
 '1-23 '0-13 '0- 20 '0-17 '0-10 '7-38 '2-17 '2-29 '0-18 '2-20
 '0-18 '2-22 '2-27 '0- 0 '0-17 '0-20 '0-17 '0-20 '1-22 '3-27

 '7-30 '0-33 '0-17 '2-00 '3-27 '0- 2 '0- 7 '0-10 '2-19 '1-29
 '7-30 '0-17 '0-17 '1-10 '7-29 '7-20 '0- 7 '1- 9 '0-17 '1-27
 '0-28 '2-22 '1-22 '7-20 '0- 8 '0-17 '0-12 '7-20 '2-27 '1-27
 '1-10 '0-18 '2-22 '2-27 '0-10 '0-18 '0-22 '7-22 '2-27 '7-23
 '0-17 '0-19 '2-21 '3-20 '0-30 '1-21 '0-12 '3-19 '2-17 '0-11

 '0-19 '2-10 '0-20 '0-27 '1-30 '1-21 '7-20 '0-19 '1-20 '0-11
 '0-20 '0-10 '1-23 '0-21 '2-19 '0-17 '0-17 '0-13 '0-20 '1-12
 '3-20 '0-17 '0-17 '2-10 '2-20 '0-19 '1-10 '2-21 '3-30 '3-17
 '2-30 '0-20 '0-19 '0-17 '1- 9 '3-22 '0-27 '2-30 '2-29 '2-19
 '1-30 '2-30 '0- 21 '1-10 '0- 0 '0- 8 '1-12 '0-20 '7-28 '9-23

 '1-10 '2-12 '0-17 '3-31 '2-19 '2-20 '1-27 '1-20 '3-30 '0-30
 '0- 0 '2-21 '0-19 '0-30 '3-20 '2-12 '0-17 '1-22 '2-22 '7-22
 '2-21 '2-20 '2-30 '0-10 '0-12 '0-17 '2-23 '0-20 '0-19 '0-19
 '2-22 '0-21 '1-20 '2-20 '0-18 '8-20 '0-23 '0-19 '7-20 '0-17
 '0-19 '1-10 '8-30 '3-28 '0-10 '0-27 '3-28 '3-32 '0-22 '0-10

عدد الأطفال	فئات الأجيال (باللبيرة)											المجموع
	٤	٨	١٢	١٦	٢٠	٢٤	٢٨	٣٢	٣٦	٤٠	٤٤	
صفر	III	IIII IIII	IIII IIII	IIII IIII	IIII IIII	IIII IIII	IIII	II	II	I	-	٧٧
١	-	IIII	IIII IIII	IIII IIII	IIII IIII	I IIII	II	IIII	-	-	-	٥٨
٢	II	-	IIII	IIII IIII	IIII IIII	IIII IIII	IIII	I	II	I	-	٤٥
٣	-	-	I	IIII	IIII	IIII	II IIII	IIII	IIII	I	-	٢٧
٤	-	-	-	I	-	-	II	IIII	II	I	I	١٥
٥ فأكثر	-	-	-	-	II	IIII	II IIII	IIII	IIII	IIII	-	٣٣
المجموع	٦	١٤	٣١	٥٤	٤٣	٣٦	٢٧	١٩	١٠	٦	٣	٢٥٠

الجنوع	فئات الأحمور (بالثيرة)												عدد الأطفال
	-٤٨	-٤٤	-٤٠	-٣٦	-٣٢	-٢٨	-٢٤	-٢٠	-١٦	-١٢	-٨	-٤	
٧٧	-	-	١	-	٢	٣	٩	١٢	١٩	١٥	١٠	٤	٦
٥٨	-	-	-	-	٣	٤	٦	١٢	٢١	١٥	٤	-	١
٤٥	-	-	١	٢	٩	٣	٩	١١	٨	٥	-	-	٢٠
٢٧	-	١	-	٢	٤	٧	٥	٤	٣	١	-	-	٣
٩٥	١	-	١	٢	٣	٥	٢	-	١	-	-	-	٤
٢٣	٣	٢	٢	٤	٦	٧	٥	٤	٢	-	-	-	٥ فاكسر
٢٥٠	١	٣	٦	١٠	١٩	٢٧	٣٦	٤٣	٥٤	٣١	١٤	٦	المجموع

وفضلاً عن تلخيص البيانات الأصلية ووضعها في شكل يسمح بتفهم العلاقات فيما بينها فإن الجدول السابق يمكننا من الوقوف على طبيعة كل من المتغيرين على حده فيمكن مثلاً معرفة التوزيع التكراري للأجور بصرف النظر عن عدد الأطفال، وكذلك يمكن الوقوف على طبيعة التوزيع التكراري لعدد الأطفال بصرف النظر عن توزيع الأجور .

ويمكن في الواقع استخلاص عدة توزيعات تكرارية بسيطة من واقع هذا التوزيع المزدوج ، فمثلاً يمكن معرفة التوزيع التكراري لأطفال العمال الذين يتقاضون من ٢٤ قرشاً إلى أقل من ٢٨ قرشاً وذلك باستخدام تكرارات العمود الأول والسابع من الجدول المزدوج ، وبالمثل يمكن استنباط التوزيع التكراري لأجور العمال الذين لديهم خمسة أطفال فأكثر وذلك باستخدام السطرين الأول والسابع من نفس الجدول .

وفضلاً عن ذلك فإن جدول التوزيع المزدوج يمكننا من الوقوف على حقيقة العلاقة بين المتغيرين ان كانت هناك ثمة علاقة ، فلو ألقينا نظرة سريعة على تكرارات خانات الجدول المزدوج لأمكننا القول بأنه كلما زاد الأجر للعامل زاد تبعاً لذلك (على العموم) عدد أطفال هذا العامل . فنرى مثلاً أن هناك ستة عمال يتقاضون أجراً من أربعة قروش إلى أقل من ثمانية قروش وليس لديهم أطفالاً وأن هناك أربعة عمال أجروا من ثمانية قروش وأقل من ١٢ قرشاً ولديهم طفل واحد . وهذا يفسر حقيقة كون الجدول خالياً من التكرارات عندما يكون عدد الأطفال كبيراً والأجور قليلة وكذلك عندما تكون الأجور كبيرة وعدد الأطفال قليلاً ، وبعبارة أخرى يمكن القول بأن هناك علاقة أو ارتباط بين الظاهرتين - الأجر وعدد الأطفال - بمشاهدة انتظام تكرارات الجدول في شكل شريط مائل من الطرف الأيمن في أعلى

الجدول إلى الطرف الأيسر في أسفله ان كانت العلاقة مضطربة أي الارتباط موجباً ، وينعكس وضع هذا الشريط ان كانت العلاقة عكسية أي الارتباط سالبا (هذا على فرض ان التوزيعين وضعاً في الجدول بشكل تصاعدي) . ولهذا السبب يطلق الاحصائيون على هذا الجدول اسم « جدول ارتباط » .

وانتظام التكرارات في هذا الشريط دليل على وجود الارتباط ، أما تحديد قيمة الارتباط نفسه فسنرجىء بحثه عندما ندرس بالتفصيل موضوع الارتباط .

الفصل الرابع

التوضيح البياني

رأينا كيف يستخدم التبويب في وضع عدد كبير من البيانات في صورة منظمة واضحة تساعد على استيعاب الموضوع الذي نقوم بدراسته . على ان هذه الصورة قد لا تكفي احيانا للتوضيح خاصة وان بعض الناس يجدون صعوبة كبيرة في ادراك مدلولات الارقام التي تعرض عليهم في جداول ، ولهذا يحاول الاحصائي توضيح البيانات بطريقة اخرى وهي الرسوم البيانية ، حيث ان كثيرا من هذه الرسوم يساعدنا احيانا في تكوين فكرة سريعة ودقيقة عن كثير من البيانات المعقدة . هذا فضلا عن ان بعض الرسوم تساعدنا في اجراء التحليل الاحصائي .

وتختلف وسائل العرض البياني باختلاف البيانات التي لدينا ، وذلك لأن الرسم البياني يجب ان يصمم بحيث يبرز الفكرة الاساسية التي يرغب مصمم الرسم في ابرازها ، ولهذا تكون اول خطوة في التوضيح البياني هي اختيار الرسم الذي يناسب بيانات معينة حتى تبرز الفكرة التي تدور حولها هذه البيانات . ونلاحظ ان طرق العرض البياني قد تطورت وتشعبت بحيث اصبحت فنا له قواعده واصوله . ولا يتسع المجال في هذا الكتاب لاستعراض الطرق المختلفة التي يمكن ان تستخدم في تجميل الرسوم البيانية حتى تزيد جاذبيتها . وبذلك سنقتصر على دراسة الطرق الرئيسية على ان يؤخذ في

الاعتبار دائماً ان الرسم البياني هو فن قبل ان يكون عملاً احصائياً ولهذا لا يجب ان نبخل على الرسم بكل ما يساعدنا على زيادة جاذبيته ووضوحه .

الرموز البيانية :

ولعل أول طريقة للتوضيح البياني هي الرموز البيانية وهي طريقة تستخدم غالباً في النشرات التي تنشرها الحكومة والهيئات الخاصة بقصد تفهيم افراد الجمهور بعض الحقائق الرقمية ، وقد ازداد استعمال هذه الطريقة في بعض الدول التي يكون فيها حقاً للجمهور أن يتعرف على الحقائق المختلفة التي تمس حياته الاقتصادية والاجتماعية . وتبعاً لهذه الطريقة تمثل الظاهرة موضوع التوضيح بصورة ترمز اليها (صورة مصنع للدلالة على المؤسسات الصناعية مثلاً) ثم بتغيير حجم هذه الصورة أو تغيير عدد وحداتها يمكن التعبير عن الاتجاه الرقمي لهذه الظاهرة . الا ان هذه الطريقة ينقصها الدقة في التوضيح خاصة إذا تذكرنا أن الصورة مهما كانت دقيقة لا يمكن أن تعبر بدقة عن التغير في الأرقام ، ولهذا فهي تناسب الأغراض التي تستخدم فيها فقط حيث يكون الهدف من استعمالها هو اعطاء فكرة عامة بقصد الدعاية عن الاتجاه الرقمي لظاهرة ما . أما بالنسبة للأغراض الأخرى التي تكون للدقة فيها أهمية كبيرة يحاول الاحصائي استخدام بعض الطرق الأخرى في التوضيح

شكل الأعمدة :

يستعمل شكل الأعمدة في توضيح قيم ظاهرة ما في عدة فترات زمنية وذلك لابرار التغير الذي حدث فيها ، وكذلك في توضيح قيم الأوجه المختلفة لظاهرة معينة لابرار المقارنة بين هذه الأوجه . على اننا نلاحظ انه عند استعمال هذه الطريقة في التوضيح البياني يجب أن لا نغالي في استعمالها حيث أن كثرة الأعمدة في الرسم تجعله يبدو مزدحماً غير مقبولا .

ويتكون شكل الأعمدة من مستطيلات ذات سمك واحد ومنفصلة عن بعضها بحيث تتناسب ارتفاعاتها مع قيم الظاهرة في الفترات الزمنية المختلفة أو مع قيم الأوجه المختلفة للظاهرة . وعندما يرسم هذا الشكل على أساس مستطيلات أفقية وليست راسية تكون أيضاً المستطيلات ذات سمك واحد ومنفصلة عن بعضها بحيث تتناسب أطوالها مع القيم المختلفة للظاهرة . وإذا كانت بعض البيانات التي تمثل الظاهرة موجبة والبعض سالبة ترسم القيم الموجبة في الاتجاه الموجب من الرسم البياني والقيم السالبة في الاتجاه السالب بحيث يكون المحور الصفري فاصلاً بينها ، وذلك مثل توضيح التغير في الميزان التجاري أو في ميزان المدفوعات أو في ميزانية إحدى المؤسسات حيث تكون القيم لهذه الظواهر في بعض السنوات موجبة وفي بعض السنوات سالبة .

ومن الواضح أنه يحسن إبراز هذه المستطيلات باللون الأسود أو بتظليلها بواسطة خطوط متوازية مرسومة داخلها ، وعندما تمثل المستطيلات فيما موجبة وأخرى سالبة يحسن تلوينها بالوان مختلفة أو تظليلها بطريقتين مختلفتين .

وفي هذا النوع من الرسومات البيانية يجب أن يبدأ المقياس المدرج من الصفر لأننا إذا بدأنا من أي قيمة أخرى فسوف تختلف النسب بين ارتفاعات الأعمدة أو أطوال المستطيلات عن النسب الحقيقية للأرقام التي تمثلها هذه المستطيلات . ولتوضيح ذلك نفترض أن القيم التي نريد إظهارها في الرسم هي ٣٠٠ ، ٢٠٠ ، ١٠٠ ، أي النسبة الحقيقية بينها هي ٣ : ٢ : ١ فإذا بدأنا بالرقم ٥٠ وليس بالرقم صفر على المقياس المدرج وجعلنا كل سم = ٥٠ وحدة - فإن ارتفاع العمود الأول يكون ٥ سم وارتفاع العمود الثاني ٣ سم وارتفاع العمود الثالث ١ سم أي تكون النسبة بين ارتفاعات المستطيلات هي : ٥ : ٣ : ١ وهي تختلف اختلافاً كبيراً عن النسبة بين الأرقام التي

توضحها هذه المستطيلات . كذلك يحسن ان تكون المسافة بين كل مستطيل وآخر أو بين كل عمود وآخر نصف المقياس الذي يدل على سمك العمود أو المستطيل . ومن الأفضل عدم كتابة الأرقام التي تمثلها الأعمدة فوق الأعمدة حيث ان ذلك يبالغ في طول الأعمدة بالإضافة إلى جعل الرسم يبدو منفراً .

ويمكن أن تكون الظاهرة ذات قيمتين أو أكثر في كل فترة من الفترات الزمنية ، وفي هذه الحالة إذا كان من المنطق جمع هاتين القيمتين بحيث يكون المجموع دالاً على القيمة الكلية للظاهرة يرسم عموداً لكل فترة ويقسم إلى الجزئين بحيث يكون الارتفاع الكلي للعمود معبراً عن القيمة الكلية للظاهرة وارتفاع كل جزء منه معبراً ومتناسباً مع قيمة كل جزء . ويمكن استعمال هذه الطريقة مهما كان عدد الأجزاء التي تتكون منها الظاهرة وسواء أردنا التوضيح باستعمال الأعمدة الرأسية أو المستطيلات الأفقية .

أما إذا لم يكن من المنطق جمع هذه الأجزاء يرسم لكل فترة عمودين متلاصقين ومتساويين في القاعدة ومختلفين في الارتفاع تبعاً لاختلاف القيم التي تمثلها هذه الأعمدة . فإذا أردنا توضيح الصادرات والواردات في عدة فترات زمنية لا يكون من المنطق جمعها في عمود واحد ، بل المنطق هو في توضيحها بعمودين متلاصقين حتى يبرز الفرق وهو الذي يدل على الميزان التجاري. ويمكن استعمال هذه الطريقة مهما كان عدد الظواهر التي نريد إظهارها في كل فترة من الفترات . وسواء أردنا التوضيح باستعمال الأعمدة الرأسية أو المستطيلات الأفقية .

ومن الواضح أنه في هاتين الحالتين لا يجب أن نكثر من الأعمدة في الرسم حتى يبدو مكتظاً بها فيضيع بذلك الهدف الأساسي من التوضيح . وبشكل عام لا يجب أن نعرض أكثر من ثلاث ظواهر عن طريق رسم أعمدة أو مستطيلات متلاصقة حيث انه كلما زاد عدد الأنواع المختلفة من الأعمدة كلما أصبح الرسم معقداً وبذلك تقل فائدته .

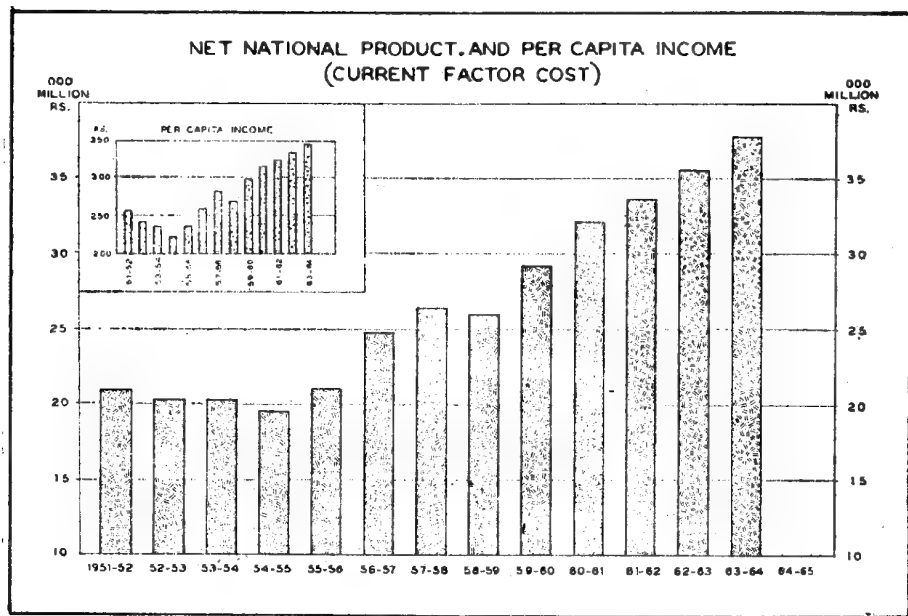
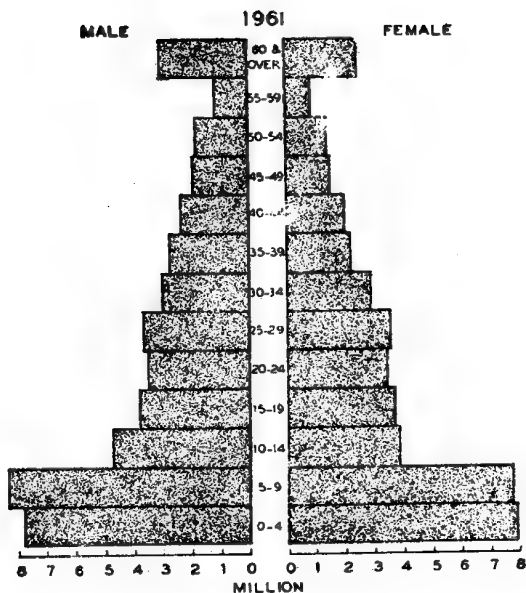
كذلك يجب التمييز بين أجزاء الأعمدة أو بين الأعمدة التي تمثل الظواهر المختلفة في كل فترة بالألوان أو التظليل ، على أنه لا يجب أن نلجأ الى الألوان إلا اذا كان لدينا الخبرة في استعمالها والا فان الرسم سوف يبدو قبيحاً منفراً.

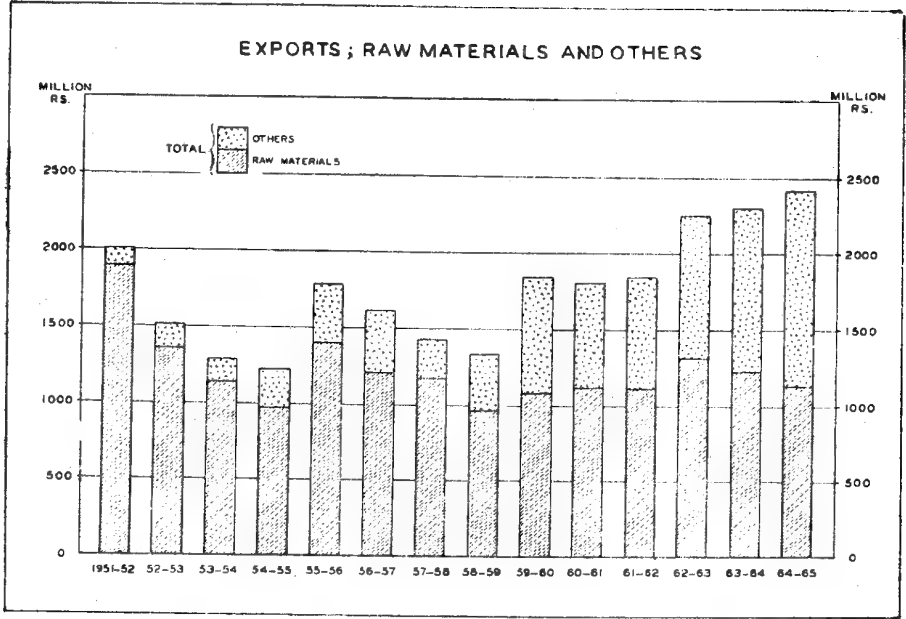
وكذلك يمكن أن يستعمل شكل الأعمدة في توضيح التغير الذي حدث في عدة ظواهر بين فترتين زمنيتين (عامين مثلاً) ، ويكون ذلك يجعل المحور الرأسي للرسم في الوسط ثم نرسم مستطيلات في الاتجاه الموجب تبين الزيادة التي حدثت في الظواهر التي ازدادت قيمها ومستطيلات أخرى في الاتجاه السالب تبين النقص الذي حدث في الظواهر التي نقصت قيمها . ويجب أن نلاحظ أن أطوال المستطيلات لا تتناسب مع قيم هذه الظواهر وانما مع مقدار الزيادة أو النقص الذي حدث فيها ومن الواضح أنه يجب التقهقرة بين المستطيلات التي في الاتجاه الموجب وتلك التي في الاتجاه السالب بالتلوين أو باستعمال التظليل المختلف .

ويمكن أيضاً استخدام شكل الأعمدة في توضيح التركيب الجنسي والعمرى للسكان في احدى الدول حيث توضح فئات الاعمار بين عمودين في منتصف الرسم ، ثم نرسم مستطيلات تبين عدد الذكور أمام كل فئة من الفئات في الجهة اليمنى من الرسم ومستطيلات تبين عدد الإناث أمام كل فئة من فئات الاعمار في الجهة اليسرى من الرسم . وفي هذه الحالة يكون المقياس المدرج على المحور الافقي وعلى أساسه تتناسب أطوال المستطيلات مع عدد الذكور وعددا الإناث في كل فئة من فئات الأعمار . ويسمى هذا الرسم بهرم التركيب العمرى للسكان وله فوائد كثيرة حيث أنه يوضح النمط السكاني الذي يمر به المجتمع ، فاذا كان نمطاً بدائياً فان قاعدة الهرم تكون واسعة جداً بالنسبة للمدرجات الاخرى التي فوقها حيث تكون معدلات المواليد في هذه الدول مرتفعة جداً ، وكذلك تقلص درجات الهرم تقلصاً سريعاً مما يدل على ارتفاع معدلات الوفاة . اما اذا كان نمطاً انتقالياً فان قاعدة الهرم تكون واسعة كذلك دلالة

على ارتفاع معدلات المواليد ، الا أن درجات الهرم لا تقلص سريعاً دلالة على انخفاض معدلات الوفيات (السكان في الدول الشرقية) . أما اذا كان نمطه يدل على النضوج السكاني فان قاعدة الهرم تكون ضيقة نسبياً دلالة على انخفاض معدلات المواليد كثيراً ، كما تكون درجات الهرم متقاربة جداً دلالة على انخفاض معدلات انوفيات الى أدناها . وبذلك يرتفع الهرم كثيراً دلالة على ارتفاع العمر المتوقع للطفل المولود . كذلك يمكن أن يدلنا الهرم على ما اذا كان السكان يتأثرون في نموم بالعوامل الطبيعية فقط (المواليد والوفيات) أو بعوامل شاذة كذلك (الهجرة والحروب .) حيث أن الدول التي يهاجر اليها السكان ينبعج الهرم الخاص بها في فئات العمر الخاصة بالشباب من ناحية الذكور والعكس اذا كان يهاجر منها السكان بتقلص الهرم تقلصاً شاذاً في هذه الفئات . كذلك يمكن أن يدلنا الهرم على أثر الحرب على النمو السكاني حيث نلاحظ ان مستطيلات الذكور تكون في الغالب أقصر من مستطيلات الاناث ، وكذلك نلاحظ تقلص المستطيلات الخاصة بالمواليد بسبب تغيب الذكور أثناء الحرب .

ونلاحظ انه في كل هذه الأنواع من الرسوم التوضيحية يمكن أن نلجأ اليها سواء كانت الارقام التي لدينا ارقاماً مطلقة أو أرقاماً نسبية ، فبالنسبة لهرم التركيب العمري مثلاً يمكن أن تدل المستطيلات على عدد كل من الذكور والاناث في كل فئة عمر ويمكن أن تدل على نسبة الذكور أو الاناث في كل فئة الى المجموع الكلي للسكان والفارق الوحيد ان المقياس المدرج في الحالة الاولى يظهر أرقاماً مطلقة أما في الحالة الثانية فيظهر أرقاماً نسبية .





21

شكل الدائرة :

يستعمل شكل الدائرة في عرض التوزيع النسبي للأوجه المختلفة لظاهرة معينة في فترة زمنية واحدة ، على أن تكرر هذه الأوجه مكونة للمجموع الكلي لهذه الظاهرة ، مثلاً توزيع المساحة الكلية للأراضي الزراعية في الدولة تبعاً لأوجه استخدامها المختلفة ، والأنواع المختلفة للقروض التي منحها أحد المصارف ... الخ . ولرسم هذا الشكل التوضيحي نرسم دائرة ذات قطر مناسب (ليس كبيراً جداً ولا صغيراً جداً) ونقسمها إلى قطاعات تتناسب مساحتها مع نسبة كل جزء إلى المجموع الكلي . ولاجراء ذلك نقسم الزاوية

المركزية والتي نسوي 360° تبعاً لعدد الاجزاء التي تتكون منها الظاهرة وتبعاً لنسبة كل جزء من المجموع . فاذا فرضنا أن مجموع الاجزاء ٧٢٠ الف دونم وقيمة الجزء الاول ٣٤٦ الف دونم وقيمة الجزء الثاني ١٠٨ الف دونم وقيمة الجزء الثالث ٣٦٦ الف دونم يكون تقسيم الزوايا كالآتي :

$$172,8^\circ = 360^\circ \times \frac{346}{720} = 360^\circ \times 0,48$$

$$54,0^\circ = 360^\circ \times \frac{108}{720} = 360^\circ \times 0,15$$

$$123,2^\circ = 360^\circ \times \frac{266}{720} = 360^\circ \times 0,27$$

وعلى أساس هذه الزوايا ١٧٢,٨ ، ٥٤ ، ١٢٣,٢ نقسم مساحة الدائرة الى القطاعات الثلاث ، ويمكن تظليل هذه القطاعات أو تلوينها للتمييز بينها ويكتب داخل كل قطاع اسم الجزء الذي يمثله . ويمكن أن يكتب على محيط الدائرة وخارج كل قطاع النسبة التي يمثلها من المجموع الكلي .

ومن الواضح انه لا يمكن استعمال شكل الدائرة في مقارنة التوزيع النسبي لظاهرة معينة في سنتين أو أكثر حيث لا يتضح من الرسم المقارنة بين قيم الاقسام المتناظرة خاصة إذا كانت قيمها متقاربة في السنتين .

السلسلة الزمنية :

سبق ان ذكرنا ان شكل الأعمدة لا يتفق مع البيانات الخاصة بعدد كبير من السنوات حيث ان الرسم يظهر مكتظاً بالأعمدة فيفقد بذلك رونقه ، لذلك يكون من الافضل في هذه الحالة رسم سلسلة زمنية خاصة اذا كانت البيانات تمثل متغيرات متصلة أي يكون تغيرها مع الزمن تغيراً متصلاً، والسلسلة

الزمنية تساعد بذلك في إظهار الاتجاه العام للظاهرة موضوع البحث والتغير الذي يحدث في هذا الاتجاه العام من فترة زمنية الى أخرى .

وفي هذا النوع من الرسم تظهر قيمة الظاهرة في كل فترة بنقطة معينة على ارتفاع يمثل هذه القيمة ، ثم نصل هذه النقاط ببعضها حسب تسلسلها الزمني . وإذا كنا نريد أن نوضح الاتجاه العام للظاهرة نرسم خطاً أو منحني يتوسط النقاط التي حددناها مقدماً خير توسط اي بحيث يكون مجموع انحرافات النقاط عنه يساوي صفراً تقريباً ، إذ أننا في توضيح الاتجاه العام لا يهمنا إظهار التغير من عام إلى آخر وإنما الشكل العام لاتجاهها سواء كان مستقيماً صاعداً أو هابطاً أو منحنياً محدباً إلى أعلى أو مقعراً إلى أسفل أو أي شكل آخر من الاشكال الرياضية التي تمثلها المعادلات المختلفة ، الدرجة الأولى أو الثانية أو الثالثة ... الخ .

ويمكن أن يحتوي الرسم على أكثر من سلسلة زمنية ، وفي هذه الحالة يجب توضيح كل منها بلون مختلف أو بأية طريقة مختلفة (خط متصل أو خط مقطع ... الخ) حتى نبرز اتجاه كل منها . ويحدث ذلك عندما نريد أن نظهر التطور الزمني في أكثر من ظاهرة واحدة ، على انه يحسن ان لا يظهر في الرسم أكثر من خمس خطوط بيانية ، إذ لو زاد العدد عن ذلك فإن تقاطعها أثناء اتجاهها يجعلها معقدة غير واضحة فتضيع الفائدة منها . وفي بعض الأحيان يكون الفراغ بين سلسلتين معبراً عن ظاهرة معينة أخرى نريد إبرازها في الرسم وبذلك نعمل على تظليلها بحيث لا يطمس التظليل معالم اتجاه السلسلتين اللتين تحددان الفراغ ، فعند رسم سلسلة زمنية للعوايد وأخرى للوفيات يكون الفراغ بينها معبراً عن الزيادة الطبيعية في السكان خلال الفترة موضوع البحث .

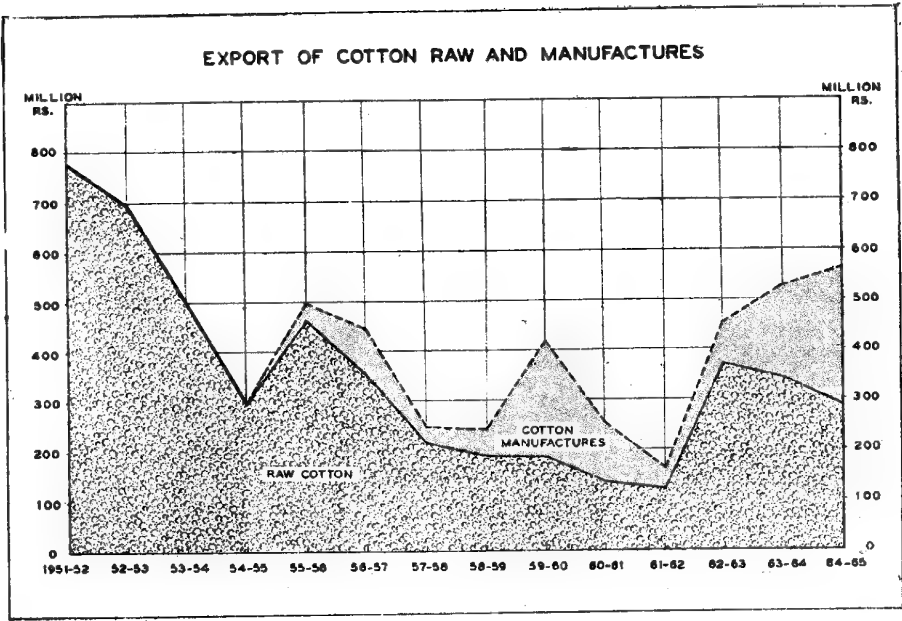
وفي بعض الأحيان يكون لدينا ظاهرة معينة مقسمة الى أجزاء في عدة سنوات ونريد إبراز التقسيم بجانب إبراز التغير الذي حدث فيه خلال

هذه المدة . لا نستطيع أن نوضح ذلك بشكل الدائرة حيث انه سبق ان ذكرنا ان التقسيم النسبي في عدة سنوات لا يمكن أن تظهره الدوائر الخاصة بالسنوات المختلفة . ومن الممكن توضيح هذه البيانات بأعمدة مجزأة للسنوات المختلفة التي يتضمنها البحث ، إلا ان كثرة الأعمدة كما ذكرنا يضيع على من يطلع على الرسم ملاحظة التغير الذي يصيب الأجزاء ، من عمود إلى آخر ، أي من فترة الى أخرى . لذلك يكون من الأفضل رسم ما نسميه بشكل القطاعات ، وهو في الواقع لا يختلف عن شكل السلسلة الزمنية إلا في حقيقة هامة وهي ان الخط البياني نفسه لا يمثل اتجاه الجزء الأول خلال سنوات البحث وإنما الذي يمثله هو الفراغ المحصور بين كل خط بياني وخط آخر ولذلك نسميه بشكل القطاعات . ولرسم هذا الشكل نرسم أولاً خطأ بيانياً يمثل اتجاه الجزء الأول من الظاهرة خلال سنوات البحث ثم خطأ بيانياً آخر يمثل اتجاه مجموع الجزئين الأول والثاني (وبذلك يكون الفراغ بينه وبين الخط الأول معبراً عن الجزء الثاني) ثم خطأً بيانياً ثالثاً يمثل اتجاه مجموع الثلاث أجزاء سوياً (وبذلك يكون الفراغ بينه وبين الخط الثاني معبراً عن اتجاه الجزء الثالث) . وهكذا حتى نصل الى آخر الخط الذي يمثل في الواقع اتجاه المجموع الكلي للظاهرة خلال سنوات البحث . بعد ذلك نظل كل قطاع أو نلونه بلون مختلف ونكتب اسمه في داخله . والرسم بهذه الطريقة يوضح التركيب الجزئي للظاهرة في كل سنة وكذلك التغير في هذا التركيب من عام الى آخر .

كذلك يمكن ان يكون لدينا القيم الخاصة بظاهرة معينة في عدة فترات زمنية ونريد ان نوضح التغير الذي حدث في هذه القيم الذي يكون في بعض الاحيان موجباً وفي أحيان أخرى سالباً (مثلاً التغير في الأرقام القياسية حيث يمكن أن يزيد الرقم عن ١٠٠ وبذلك يكون هناك تغير موجب ويمكن أن تنقص القيمة عن ١٠٠ وبذلك يكون هناك تغير سالب) . والرسم الذي يوضح هذه الحالة لا يختلف كثيراً عن الرسم العادي للسلسلة الزمنية والفرق الوحيد

بينهما هو في تقسيم الرسم الى جزئين يوضح الجزء الأعلى منه التغيرات الموجبة والجزء الأسفل يوضح التغيرات السالبة . ثم نحدد موضع النقط في الرسم حسب قيمة التغير وحسب إذا كان موجباً أو سالباً ثم نوصل النقط ببعضها حسب تسلسلها الزمني من فترة الى أخرى

ويمكن أن يكون لدينا قيم خاصة بعدة ظواهر في عدة فترات زمنية بحيث تكون وحدة القياس لهذه القيم مختلفة ، كما تختلف الأرقام عن بعضها اختلافاً كبيراً حيث يكون بعضها صغيراً جداً والبعض الآخر كبيراً جداً ، وبذلك نقع في مشكلة تحديد مقياس الرسم الذي يتفق مع هذا النوع من الأرقام . والمثل على ذلك بيانات عن عدد المؤسسات وعدد المشتغلين فيها ومتوسط الأجر للمشتغل الواحد في عدة سنوات ، ونريد أن نوضحها بحيث يمكن أن يظهر في الرسم مقارنة التغير الذي حدث فيها خلال المدة موضوع البحث ، والطريقة لمعالجة هذه الحالة هي في تحويل الأرقام المطلقة إلى أرقام نسبية على أساس الفترة الأولى وبذلك تصبح وحدة الأرقام متشابهة (%) ، كما تصبح متقاربة من بعضها . ثم ننفذ الرسم على أساس هذه الأرقام النسبية وسوف نلاحظ من الرسم انها جميعاً تبدأ من نقطة واحدة (١٠٠) ثم تتخذ اتجاهات يتفق مع التغير الذي حدث فيها وبذلك تظهر المقارنة فيما بينها ويجب ان نلاحظ كتابة عنوان الرسم كالآتي - التغير النسبي في عدد المؤسسات والمشتغلين فيها ومتوسط الأجر في المدة من عام - الى عام - .



14

خرائط الدبابيس :

تستخدم هذه الخرائط لتوضيح البيانات التي تطرأ عليها تغيرات من وقت إلى آخر ، مثل انتشار وباء معين خلال فترة زمنية معينة وانتشار الجراد وتحركات الجيوش . وفي هذا النوع من الرسم يمكن أن يعبر الدبوس الواحد عن وحدة وحدة أو عن عدد معين من الوحدات ، وتوضع هذه الدبابيس على الخرائط حسب المكان الذي ظهرت فيه هذه الحالة أو الحالات ، وهكذا كلما ظهرت حالة أو عدد معين من الحالات يوضع الدبوس على الخريطة في المكان المعين ، وبذلك يستطيع المعينون بالامر تتبع ما يحدث في الظاهرة من وقت إلى آخر . ويمكن استعمال دبابيس بالوان مختلفة أو بأشكال هندسية مختلفة لتوضيح ظواهر مختلفة على الخريطة . ومن الواضح أن مثل هذه الخرائط لا يقصد منها العرض في النشرات الاحصائية وانما العرض المكتبي لاطلاع المسؤولين .

كذلك يمكن استخدام الخرائط المنقطة لتوضيح البيانات الخاصة بظاهرة معينة في أجزاء الدولة المختلفة ، حيث توضع نقط مستديرة تمثل كل منها عدداً معيناً من البيان الخاص بالظاهرة موضوع العرض ، فلو رسمنا خريطة للبنان وقسمناها إلى محافظات فانه يمكن التعبير مثلاً عن انتاج اللواكه في كل محافظة بمثل هذا النوع من الخرائط . هذه الخرائط صالحة للطبع في نشرات حيث انها ليست مثل خرائط الدبايس إذ المقصود منها اظهار توزيع الظاهرة على أجزاء الدولة في وقت معين وليس إظهار التغير الذي يحدث في انتشار الظاهرة من وقت الى آخر .

الرسم البياني اللوغاريتمي :

بحسب الاحتياجي أحياناً الى توضيح معدل التغير في الظواهر وليس التغير المطلق في قيمها . ولتفسير ذلك نفترض ان لدينا عدد السكان في مدينتين في تعدادين ونريد أن نقارن التغير في هذه الأعداد لنعرف أي المدينتين تنمو بسرعة أكبر . نستطيع أن نحدد مقدار التغير في عدد سكان كل مدينة بطرح عدد سكانها في التعدادين ثم نقارن ناتج الطرح ، غير انه لكي تكون المقارنة سليمة لا بد أن ننسب عدد سكان كل مدينة الى عددها في التعداد الأول حيث ان ذلك هو الذي يوصلنا فعلاً الى المقارنة السليمة بينها حيث أن هدفنا هو في الواقع معرفة أي المدينتين تنمو بمعدل أكبر ؟ فإذا فرضنا ان عدد سكان المدينة الأولى هو ١٠ ملايين نسمة في تعداد ١٩٥٠ وأصبح ١٢ مليون نسمة في تعداد ١٩٦٠ ، وان عدد السكان المدينة (ب) هو ٥ مليون نسمة في تعداد ١٩٥٠ وأصبح ٧ مليون نسمة في تعداد ١٩٦٠ ، هل نستطيع أن نسقنتج من ذلك ان معدل النمو في سكان المدينتين متساو ؟ ان هذه النتيجة تكون في الواقع نتيجة مضللة حيث ان سكان المدينة (أ) قد زادوا بنسبة ٢٠٪ بينما زاد سكان المدينة (ب) بنسبة ٤٠٪ ، لذلك لا نستخدم الرسم البياني العادي في توضيح معدلات التغير حيث أن

هذا الرسم لا يستطيع أن يظهر الا التغيرات المطلقة والرسم البياني اللوغاريتمي هو الوسيلة لتوضيح معدلات التغير حيث أن الفكرة الأساسية لهذا الرسم هي أنه بينما تكون المسافات المتساوية في الرسم البياني العادي ممثلة لقيم مطلقة متساوية تكون في الرسم البياني اللوغاريتمي ممثلة لمعدلات متساوية ولذلك نجد ان مقياس التدرج في الرسم البياني العادي يظهر الأرقام على أساس أنها تكون متوالية عديدة مثلا صفر ، ٥٠ ، ١٠٠ ، ١٥٠ ، ٢٠٠ ، وهكذا ، أما في الرسم البياني اللوغاريتمي فتظهر الأرقام على أساس أنها تكون متوالية هندسية ، ١٠٠ ، ١٠٠٠ ، ١٠٠٠٠ ، وهكذا وبذلك بينما يكون كل ١ سم في الرسم البياني العادي ممثلا لخمس وحدة مثلا ، يكون كل ١ سم في الرسم البياني اللوغاريتمي ممثلا لعشرة أضعاف .

تقوم فكرة الرسم البياني اللوغاريتمي على أساس الحقيقة الهامة وهي انه اذا كان لدينا ارقام تكون متوالية هندسية فان لوغاريتمات هذه الأرقام تكون متوالية عددية أى يكون الفرق المطلق بين أعدادها متساويا مثلا :

الاعداد	١٠٠	١٥٠	٢٢٥	٣٢٧,٥	٥٠٦,٢٥
لوغاريتماتها	٢	٢,١٧٦	٢,٣٥٢	٢,٥١٨	٢,٧٠٤

ولاثبات ذلك نفترض ان لدينا القيم $س١$ ، $س٢$ ، $س٣$ ، $س٤$ ، تكون متوالية هندسية :

$$\frac{س٢}{س١} = \frac{س٣}{س٢} = \frac{س٤}{س٣}$$

$$لو \frac{س٢}{س١} = لو \frac{س٣}{س٢} = لو \frac{س٤}{س٣}$$

لو $س٢$ - لو $س١$ = لو $س٣$ - لو $س٢$ = لو $س٤$ - لو $س٣$
 أي أن لوغاريتمات $س١$ ، $س٢$ ، $س٣$ ، $س٤$ (وهي تكون هندسية كما قدمنا) تكون متوالية عددية .

وبذلك يمكننا ان نستفيد من هذه القاعدة في تصميم رسم بياني يوضح معدلات التغير في البيانات التي لدينا ، وذلك بدلا من ان نظهر القيم الفعلية لهذه البيانات (وهو يظهر تغيرها المطلق) نظهر لوغاريتمات هذه الاعداد وبذلك يتضح تغيرها النسبي ، ومن الواضح ان الظاهرة اذا كانت تتغير بمعدل ثابت فان لوغاريتمات القيم الخاصة بها سوف تتغير بمقدار ثابت ، واذا كانت قيم الظاهرة تتغير بمعدلات متزايدة كان لوغاريتماتها تتغير بمقادير متزايدة ، والعكس اذا كانت قيم الظاهرة تتغير بمعدلات متناقصة فان لوغاريتماتها تتغير بمقادير متناقصة . وبذلك يتضح لنا انه يمكن ان نوضح معدل التغير في ظاهرة ما على القياس العادي بشرط الا نرصد قيم الظاهرة نفسها بل نرصد لوغاريتمات هذه القيم . ولتنفيذ الرسم بهذه الطريقة نضع القيم على المحور الرأسي على ابعاد تتناسب مع الفرق بين لوغاريتماتها . وعند تقسيم المحور الرأسي تقسيما لوغاريتميا نضع على نقطة الأصل أي قيمة موجبة بخلاف الصفر (إذا وضعنا الرقم ١ يكون لوغاريتمه = صفر) ، ثم نأخذ على المحور بعداً يساوي لوغاريتم ٢ ونكتب عليه ٢ . حيث أن لوغاريتم ٢ = ٣.٠١ ، فانتنا نكتب القيمة ٢ على بعد ٣.٠١ ، سم من نقطة الأصل ، إلا أن ذلك يجعل الرسم من الصفر بحيث يتعذر تحديد النقط عليه فيما بعد . لذلك يكون من الأفضل أن نأخذ مسافة ٣.٠١ ، سنتمتر مثلاً لتعبر عن القيمة ٢ ثم نأخذ من نقطة الأصل مسافة تساوي لوغاريتم ٣ بنفس الوحدات التي استعملناها للتعبير عن لوغاريتم ٢ ، ونكتب على هذه المسافة الثانية الرقم ٣ وهكذا بالنسبة لباقي القيم .

بعد ذلك نستطيع أن نضع أمام التقسيمات التي قسمناها أي سلة من الأرقام تكون متناسبة مع الأعداد الطبيعية من ١ إلى ١٠ فنستطيع مثلاً أن نبدأ بالعدد ١٠٠ بدلاً من ١ بحيث نضع ٢٠٠ في مكان ٢ و ٣٠٠ في مكان ٣ وهكذا إلى أن نصل إلى الف ، كذلك يمكن أن نبدأ بالقيمة ١٥ فنكون باقي القيم ٣٠ ، ٤٥ ، ٦٠ وهكذا حتى نصل إلى ١٥٠ ، كذلك يمكن أن

نبدأ بالقيمة ٥٠ فيكون باقي القيم ١٠٠ ، ١٥٠ ، ٢٠٠ ، ٢٥٠ ، ٣٠٠ . حتى
نصل إلى ٥٠٠ . ومن الواضح اننا نختار القيم التي تتفق مع القيم التي نريد
توضيح معدل تغيرها بالرسم البياني . والمهم أن نلاحظ انه مهما غيرنا من القيم
فلا بد أن تتناسب لوغاريتماتها مع لوغاريتمات الاعداد الطبيعية من ١ الى ١٠ .
ويتضح السبب في ذلك مما يأتي : —

$$\text{لو } ١٠٠ - \text{لو } ٥٠ = \frac{١٠٠}{٥٠} = \text{لو } ٢ - \text{لو } ١$$

$$\text{لو } ٣٦ - \text{لو } ١٨ = \frac{٣٦}{١٨} = \text{لو } ٢ - \text{لو } ١$$

$$\text{لو } ٩٠ - \text{لو } ٤٥ = \frac{٩٠}{٤٥} = \text{لو } ٢ - \text{لو } ١$$

وبذلك تكون المسافة بين لو ١٥٠٠ ولو ٧٥٠ هي نفس المسافة بين
لو ٤٨ ولو ٢٤ هي نفس بين لو ٢ ولو ١ .

بعد تحديد المقياس اللوغاريتمي على المحور الرأسي نستخرج لوغاريتمات
القيم التي نريد توضيحها ونرصدها في الرسم تبعاً للمقياس الذي حددناه . وحتى
تتفق اللوغاريتمات مع لوغاريتمات الأعداد الطبيعية نقسم القيمة التي نريد رصدها
في الرسم على أول قيمة بدأنا بها على المحور الرأسي ثم نستخرج لوغاريتم ناتج
القسمة ، فإذا كنا قد بدأنا المحور بالقيمة ٥٠ ونريد رصد القيمة ٥٥ على
أساس المقياس اللوغاريتمي ، نقسم أولاً ٥٥ على ٥٠ = ١,١ ثم نستخرج
لوغاريتم هذه القيمة ونرصده تبعاً للمقياس الذي اتخذناه على المحور الرأسي
معبراً عن اللوغاريتمات .

لاحظنا أن القيم التي يمكن رصدها تبعاً للمقياس اللوغاريتمي الذي شرحناه
تتغير بنسبة تتراوح من ١ إلى ١٠ ، مثلاً ٥٠ الى ٥٠٠ ، أو ٧٠ إلى ٧٠٠ .
هذا المدى يسمى دورة لوغاريتمية . لكن ما العمل إذا كانت القيم تتغير في

مدى أوسع من ذلك مثلاً من ٥٠ إلى ٢٠٠٠ مثلاً ؟ نستطيع أن نكمل التقسيم بدورته ثانية ودورة ثالثة وهكذا - تكون الدورة الأولى من ٥٠ إلى ٥٠٠ والدورة الثانية من القيمة ٥٠٠ وتنتهي بالقيمة ٥٠٠٠ وتبدأ الثالثة بالقيمة ٥٠٠٠ وتنتهي بالقيمة ٥٠٠٠٠ وهكذا . ومن الواضح أن كل دورة يتفق تقسيمها مع تقسيم لوغاريتمات الأعداد الطبيعية من ١ إلى ١٠ . كذلك يتضح لنا أن الرسم اللوغاريتمي يمكن أن يتسع لقيم تبدأ صغيرة جداً وتنتهي كبيرة جداً .

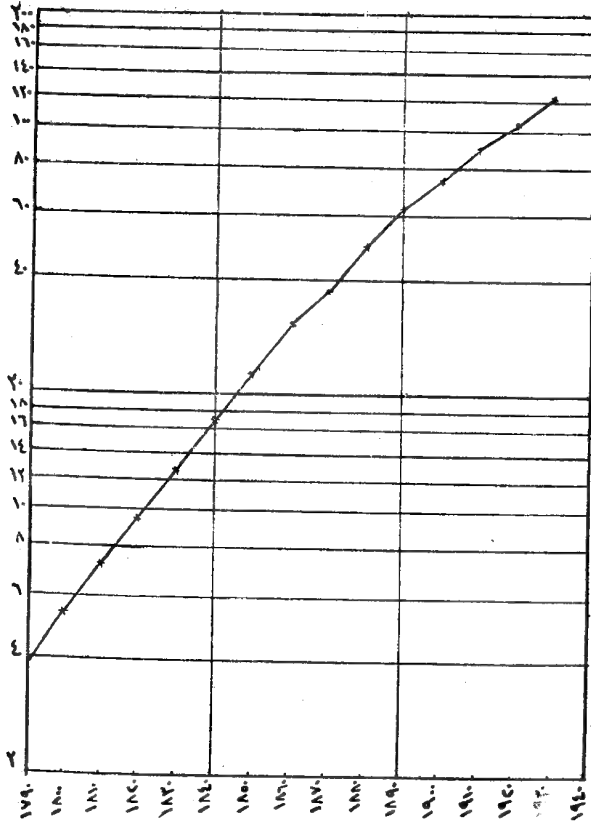
عندما يكون التقسيم اللوغاريتمي على المحور الرأسي فقط أي أن المحور الأفقي يكون تقسيمه عادياً يسمى الرسم نصف لوغاريتمي ، أما إذا كان التقسيم اللوغاريتمي على كلا المحورين يكون الرسم لوغاريتمي كامل .

بعد رصد النقط على الرسم البياني النصف اللوغاريتمي وتوصلها سوياً بخط بياني نستطيع أن نتعرف على نوع التغير في الظاهرة من ميل هذا الخط ، فإذا كانت الظاهرة تتغير بمعدل ثابت يظهر ذلك في الرسم من وقوع جميع النقط على الخط المستقيم الذي يصل بينها كما أن درجة ميل هذا الخط توضح لنا ما إذا كان التغير بمعدل متزايد أو معدل متناقص . كذلك إذا رسمنا خطين بيانيين لظاهرتين مختلفتين على رسم نصف لوغاريتمي نستطيع أن نقارن معدل التغير فيها تبعاً لميل الخطين ، فإذا توازى الخطان يدل ذلك على تساوي معدل التغير في الظاهرتين ، أما إذا اختلف ميلها يدل ذلك على اختلاف معدل التغير فيها .

وهناك تطبيقات عملية كثيرة للرسم البياني النصف اللوغاريتمي حيث يمكن أن نلجأ إليه عندما تكون الظاهرتان مقاستان بوحدات مختلفة أو عندما تكون قيمهما متباعدة كثيراً حيث تكون قيم أحدهما صغيرة جداً بينما تكون قيم الأخرى كبيرة جداً . كذلك يلجأ إلى الرسم اللوغاريتمي عندما تتغير قيم الظاهرة في مدى واسع جداً حيث تبدأ صغيرة وتنتهي بقيم كبيرة جداً . وفي المجال الاقتصادي يستخدم الرسم اللوغاريتمي لحساب مرونة الطلب حيث نرسم منحنى الطلب على رسم لوغاريتمي كامل ثم نرسم مماساً لمنحنى

الطلب عند النقطة المقابلة للثمن الذي نريد تقدير المرونة عنده ثم نحسب ميل هذا المماس على محور الثمن حيث أن هذا الميل سوف يوضح لنا ناتج قسمة التغير النسبي في الطلب على التغير النسبي في الثمن . كذلك نستخدم الرسم اللوغاريتمي في توضيح قانون باريتو الذي ينص على انه كلما زادت قيمة الدخل بنسبة معينة كلما نقص عدد الأشخاص الذين يحصلون على هذا الدخل أو أكثر بنفس النسبة، وقد أثبتت الاحصاءات أن توزيع الدخل في بعض الدول يتمشى مع هذا القانون . فاذا أخذ مقياس لوغاريتمي للدخول على المحور الافقي ومقياس لوغاريتمي آخر على المحور الرأسي لتوضيح عدد الأشخاص الذين يحصلون على هذه الدخول ورصدنا النقط التي تمثل قيم الدخول والأشخاص الذين يحصلون عليه نجد انها تقع تقريباً على خط مستقيم، الأمر الذي يدل على ان نسبة التغير في الدخل تساوي نسبة التغير في عدد الأشخاص .

كذلك نستخدم الرسم النصف اللوغاريتمي في توضيح معدل التغير في الظواهر مع الزمن مثل نمو السكان حيث أن ميل الخط الذي يمثل هذا النمو يدل على ما اذا كان السكان يزدون بمعدل ثابت أو بمعدل متزايد عندما يكون الخط في شكل منحني يصعد بسرعة من اليسار الى اليمين أو بمعدل متناقص عندما يكون الخط في شكل منحني يصعد ببطء .



التمثيل البياني للتوزيعات التكرارية :

هناك أربعة وسائل أساسية لتمثيل أي توزيع تكراري في شكل بياني :

(١) طريقة الأشرطة البيانية (الأعمدة) ، وتتميز هذه الطريقة بصلاحياتها بنوع خاص لتمثيل التوزيعات التكرارية غير الرقمية وكذلك التوزيعات ذات القيم الوتابة (الغير متصلة) ، وتقوم هذه الطريقة على أساس تمثيل المتغير على المحور الأفقي والتكرارات على المحور الرأسي فيكون الشكل البياني عبارة عن مجموعة من الأشرطة يمثل قاعدة كل منها المتغير وارتفاعه أو طوله تكرار هذه المتغير.

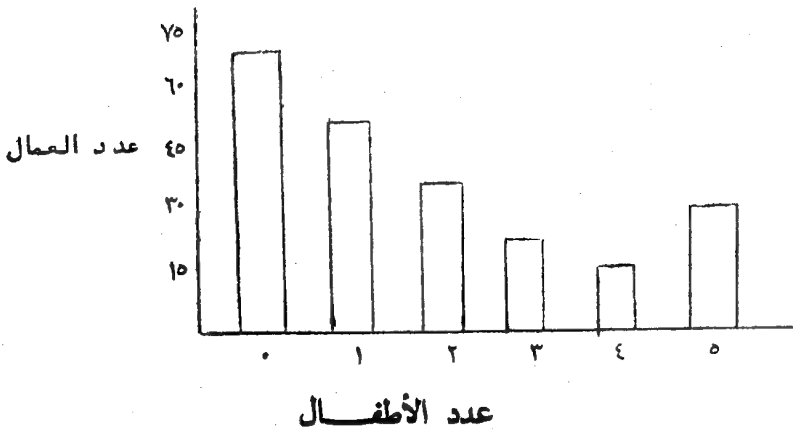
خذ التوزيع التكراري غير الرقمي المشار اليه سابقاً والخاص بتوزيع العمال حسب حالتهم الزوجية فانه من الممكن تخصيص مسافات متساوية في العرض تبعد كل منها عن الأخرى بإبعاد متساوية ثم إقامة أعمدة أو مستطيلات على كل من هذه المسافات تتناسب أطوالها وتكرار كل مجموعة كما هو مبين في الشكل التالي:



ولتوضيح تطبيق تلك الطريقة على التوزيعات التكرارية ذات المتغيرات الوثابة نأخذ تقسيم هؤلاء العمال حسب عدد أولاد كل منهم فنحصل على الجدول الآتي :

عدد الأطفال	٠	١	٢	٣	٤	٥ وأكثر	المجموع
عدد العمال	٧٧	٥٨	٤٠	٢٧	١٥	٣٣	٢٥٠

ويكون الشكل البياني كآآتي :



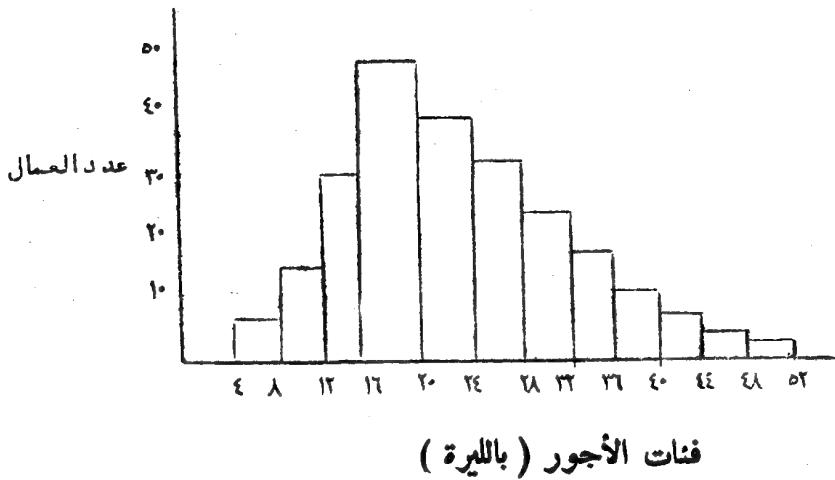
٢- والطريقة الثانية وهي الأكثر شيوعاً - هي طريقة المدرج التكراري أو الهستوجرام (Histogram) وتنحصر تلك الطريقة في أخذ مسافات على المحور الأفقي تمثل أطوال فئات التوزيع وإقامة مستطيلات ملتصقة تتناسب أطوال كل منها وتكرار كل فئة .

فلو أردنا تمثيل التوزيع التكراري للعمال السابقين حسب أجورهم فما علينا إلا ان نأخذ مسافات متساوية تمثل كل منها أربعة ليرات على المحور الأفقي وإقامة مستطيلات على تلك المسافات تمثل مساحة كل منها عدد العمال ذوي الأجر المحدد في الفئة . ولما كانت قواعد المستطيلات كلها متساوية (الفئات المتساوية) ، لهذا نكتفي برسم مستطيلات يمثل طول كل منها التكرار المقابل للفئة ، فمثلاً إذا كان التوزيع كالاتي :

فئات الأجر ٤-٨-١٢-١٦-٢٠-٢٤-٢٨-٣٢-٣٦-٤٠-٤٤-٤٨-٥٢

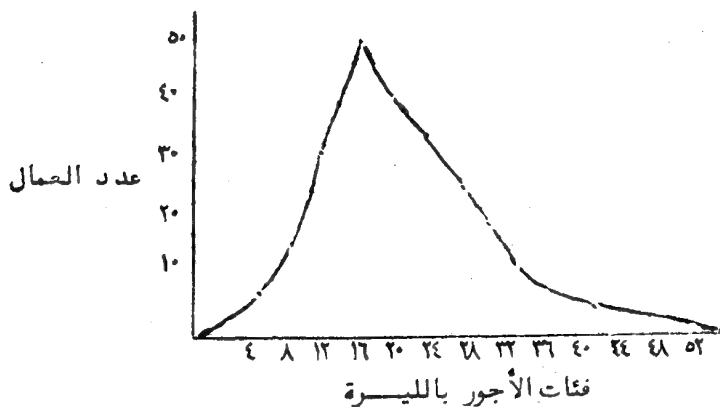
عدد العمال ٦ ١٤ ٣١ ٥٤ ٤٣ ٣٦ ٢٧ ١٩ ١٠ ٦ ٣ ١

يكون المدرج التكراري كالاتي :



ونلاحظ ان المساحة المحدودة بالمدرج التكراري هي مجموع التكرارات وذلك لأن مساحة كل مستطيل صغير في المدرج تساوي تكرار الفئة فاذا جمعنا هذه المستطيلات الصغيرة حصلنا على مجموع يمثل مجموع التكرارات هذا على أساس أن قواعد المستطيلات متساوية ولذلك يكون التعبير عنها بالرقم ١ .

(٣) والطريقة الثالثة تسمى طريقة المضلع التكراري أو البوليجون (Polygon) وتنحصر تلك الطريقة في تحديد مسافات على المحور الأفقي يمثل كل منها فئات التوزيع المطلوب رسمه ثم وضع نقطة بأعلى مركز كل فئة بحيث يتناسب بعد النقطة عن مركز الفئة مع تكرارها ثم توصيل تلك النقاط بخطوط متكسرة . والبوليجون بعبارة أخرى ما هو إلا مجموعة الخطوط المتكسرة التي توصل منتصف قمم المدرج التكراري ، وتطبيقاً لتلك الطريقة يمكن تمثيل التوزيع السابق كالآتي :



ولكي نقفل المدرج نفترض وجود فئة سابقة للتوزيع وتكرارها يساوي صفر وكذلك فئة لاحقة وتكرارها أيضاً يساوي صفر ثم نصل نهايتي المضلع بمركزي الفئتين . ونلاحظ ان المساحة المحدودة بالمضلع التكراري

تساوي المساحة المحدودة بالمدرج أي تساوي مجموع التكرارات وذلك لأن المضمع يضيف أجزاء الى مساحة المدرج ويستبعد أجزاء منها ، ومجموع الأجزاء المضافة يساوي الأجزاء المستبعدة وبذلك تبقى مساحة المدرج دون تغيير .

(٤) والطريقة الرابعة هي طريقة المنحنى التكراري Frequency Curve وهي من أهم الطرق من الوجهة النظرية .

وتتلخص هذه الطريقة في تحديد أطوال الفئات المختلفة على المحور الأفقي ثم وضع نقط بأعلى مركز كل فئة بحيث يتناسب البعد بينها وبين المركز مع تكرار الفئة ثم توصل النقط بخط ممد . وما المنحنى بعبارة أخرى الا المضمع التكراري نفسه ممدأ بحيث يتخذ شكل منحنى .

والمساحة التي يحدها المنحنى لا تساوي مساحة المدرج لأن المنحنى يضيف أجزاء ويستبعد أجزاء أخرى لا تساوها .

ويحذر بنا الاشارة الى العلاقة بين المنحنيات والمدرجات التكرارية إذ ان المنحنى التكراري يعطي في الواقع الصورة العامة للعلاقة بين المتغير وتكراراته لا للتوزيع التكراري موضوع الدراسة فحسب بل للتوزيع التكراري العام الذي اشتق منه هذا التوزيع .

خذا المثل الآتي لتوضيح تلك الفكرة :

لو أخذنا مجموعة صغيرة من طلبة الجامعة العربية ببيروت وليكن عددهم ٢٠٠ طالباً كعينة ممثلة لأفراد المجتمع الأصلي وهو طلبة الجامعة بأجمعها : ثم سجلنا أعمار أفراد تلك العينة وعلنا توزيعاً تكرارياً لهذه الأعمار فان المدرج التكراري لهذا التوزيع يعطينا العلاقة بين السن والعدد أو بعبارة أخرى يعطينا قانون تغير السن للمجموعة . تصور اننا كبرنا عدد أفراد تلك العينة وجعلناها ١٠٠٠ طالب لتمثيل المجتمع الأصلي نفسه ثم علنا توزيعاً تكرارياً لأعمار الألف طالب ورسمنا المدرج التكراري لهذا التوزيع .

واضح أن هذا المدرج التكراري سيعطينا نفس قانون التغير السابق على فرض ان كل من المجموعتين الصغيرة والكبيرة تمثل المجتمع الأصلي تمثيلاً صادقاً. تصور خطوة أخرى اننا أخذنا جميع طلبة الجامعة بلا استثناء ورسمنا مدرجاً تكرارياً لتوزيع أعمارهم فان هذا المدرج سيعطينا نفس قانون التغير السابق .

ما الفرق إذن بين المدرجات التكرارية الثلاث التي رسمناها ؟ لا فرق بينها إلا في كثرة تكرارات التوزيع الأخير بدرجة تسمح بطبيعة الحال من تقصير مدى الفئات حتى يمكن تصور ان النقط الممثلة لقمم المستطيلات المختلفة تتقارب من بعضها البعض حتى تكون خطأ مهدداً هو المنحنى التكراري الذي كان يمكن الحصول على شكله للعام بالتقريب من المدرج التكراري الأول .

نخرج من ذلك بحقيقة هامة ان المدرج التكراري يظهر قانون التغير بين أفراد عينة من مجتمع أصل بينما يظهر المنحنى التكراري قانون التغير لهذا المجتمع نفسه .

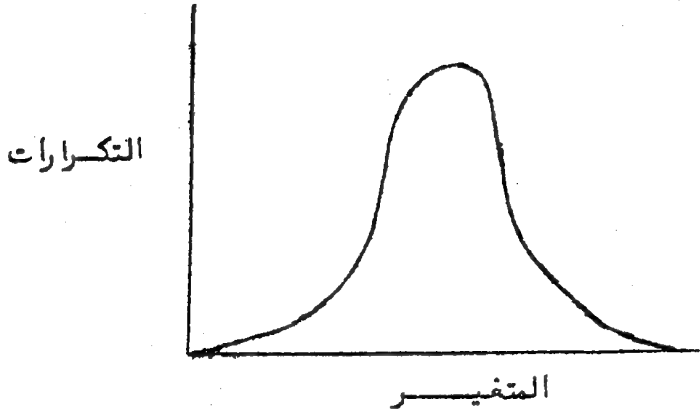
أشكال المنحنيات التكرارية :

خلصنا من دراستنا السابقة بأن شكل المنحنى التكراري لأي توزيع يعطي فكرة صادقة عن قانون التغير فيه .

ومن المعروف ان التوزيعات التكرارية مهما اختلفت في طبيعتها فانها تنقسم إلى أنواع محدودة العدد واضحة المعالم ، إذ يمكن تقسيم جميع التوزيعات التكرارية وبالتالي منحنياتها الى قسمين رئيسيين الأول - توزيعات متماثلة (symmetric) والثاني توزيعات غير متماثلة أو ملتوية (Asymmetric skewed)

والتوزيعات أو المنحنيات المتماثلة هي التي تكون العلاقة بين المتغير والتكرارات بحيث تزايد التكرارات تدريجياً كلما زاد المتغير حتى تصل التكرارات الى قمة معينة تبدأ بعدها في النزول كلما زادت القيم أكثر من ذلك ، وبحيث يكون

الصعود إلى القمة بنفس السرعة التي تهبط بها التكرارات بعد القمة كما هو واضح في الشكل التالي. وواضح ان المنحنى متماثل في صعوده وهبوطه حول قيمة معينة للمتغير ومن هنا جاءت التسمية (التماثل) .



ومذا المنحنى في شكله العام يشبه الناقوس ولذا يسميه البعض منحنى ناقوسي (Bell shaped Curve). وهذا النوع من التوزيعات من أهم الأنواع وخصوصاً في الدراسات الاحصائية النظرية . ويغلب حدوثه في التوزيعات التكرارية « الطبيعية » أي التي لا يؤثر العامل البشري على قانون التغير فيها ولذا يسميه البعض الآخر منحنى طبيعي (Normal) مثل توزيعات علم الاحياء وكذلك في علم الأجناس وفي المسائل التي يؤثر عامل الحظ (الصدفة) وحده في حدوثها مثل التوزيعات التكرارية لعدد مرات ظهور وجه معين للعملة عند رمي قطعة منها عدة مرات .

ولهذا المنحنى خواص احصائية ورياضية معينة سنحاول دراستها في الفصول القادمة .

وهناك نوع آخر من المنحنيات المتماثلة وهو المنحنى النوني (U. shaped Curvs) وهذا النوع على عكس السابق في أن أكبر التكرارات

تقابل أصغر القيم وأكبرها ، وأصغر التكرارات تقابل تقاطع خط التآثل مع محور المتغير .

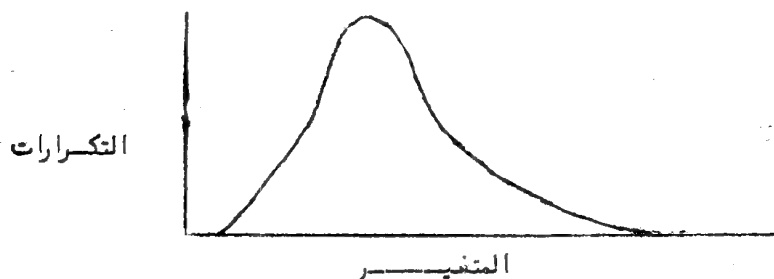
وأمثال هذه التكرارات قليلة الحدوث جداً والمثل الذي كثر ترديده بين الكتاب لهذا النوع من المنحنيات هو توزيع درجة تكاثف السحب في سماء جرينتش في سنة من السنين . ومن أمثلة هذا النوع من التوزيع التكراري نسبة البطالة بين عدد كبير من العمال بحسب أعمارهم اذ أن نسبة البطالة بين صغيري السن كبيرة ثم تتضاءل كلما كبر السن الى حد معين تبلغ عنده نهايتها الصغرى ثم ترتفع تدريجياً بعد ذلك كلما ارتفع سن العامل

ولا بد لنا أن نذكر بصراحة انه قل أن يتوفر في التوزيعات التكرارية العملية توزيعاً من هذا النوع يكون تاماً في تماثله ولكنه على أي حال يمكن اعتبار الكثير من تلك التوزيعات قريبة من التآثل بشيء من التجاوز .

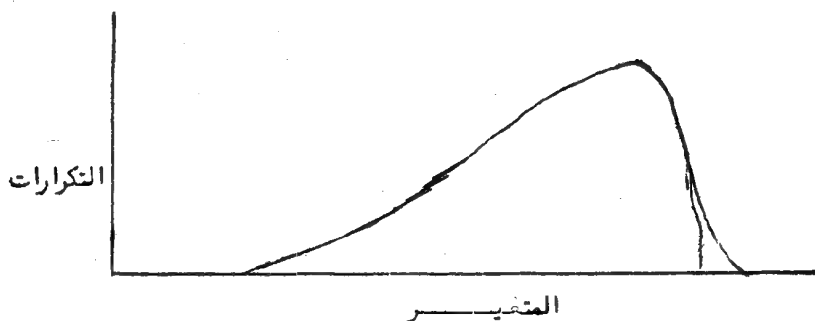
التوزيعات الملتوية :

أما المنحنيات الملتوية فهي اما معتدلة الالتواء أو حادة الالتواء . وهناك طائفتان من النوع الأول ، الأولى هي المنحنيات الموجبة الالتواء وهي ما تزايد التكرار فيها بسرعة كلما تزايدت قيمة المتغير حتى تصل الى قمتها ثم تنخفض التكرارات ببطء كلما تزايدت القيم بعد ذلك . ويستطيع القارئ أن يرى مثلاً من أمثال تلك التوزيعات في الشكل التالي وتسهيلاً له الالتواء تحديد طبيعة هذا النوع من المنحنيات يجب أن يذكر أن المنحنى الموجب الالتواء هو ما كان « ذيله » الى اليمين . ويكثر حدوث مثل هذا التوزيع في المسائل الاجتماعية والاقتصادية . حيث تميل أغلب التكرارات الكبيرة الى ناحية القيم الصغيرة للمتغير . فمثلاً لو أخذنا التوزيع التكراري للمتزوجين في بلد معين بحسب السن فان عددهم عند العمر ١٨ يكون قليلاً ثم يكبر عددهم كلما زاد العمر حتى يصل العدد أقصاه عند العمر ٢٢ ، ٢٣ سنة مثلاً ثم يبدأ

عدهم يتناقص لكل فئة من العمر ببطء حتى يصل العدد الى أقله عند العمر ٦٠ ، ٧٠ سنة .

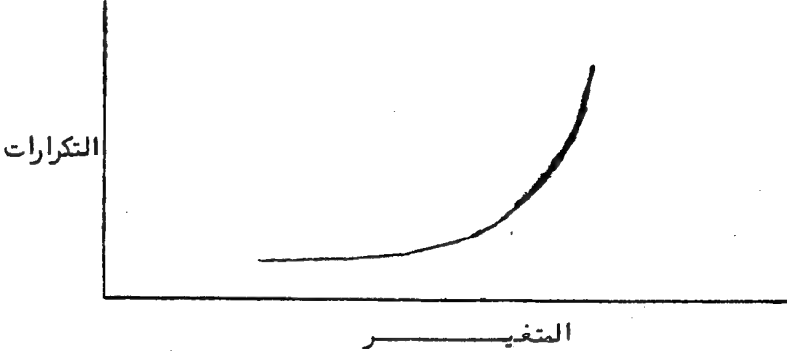


والطائفة الثانية من المنحنيات المعتدلة الالتواء هي التي تمتاز ببطء صعود تكراراتها كلما زاد المتغير حتى تصل التكرارات الى قمته عند نقط معينة ثم تبدأ في النزول بسرعة كلما زاد المتغير بعد ذلك ويسمى هذا النوع منحني ملتو التواء سالباً حيث يكون « ذيله » الى اليسار كما هو واضح في الشكل التالي . وهذا النوع من المنحنيات قل أن يصادفه الباحث وخصوصاً بين الظواهر الاجتماعية والاقتصادية ولذلك فأمثلة في هذا الميدان قليلة ، وعلى الرغم من ذلك اذا عملنا توزيعاً تكرارياً لأعمار المواليد بحسب أعمار الأمهات فان التوزيع التكراري الناتج يكون من هذا النوع . ويكثر على العكس من ذلك أمثال هذا التوزيع في المسائل الخاصة ببعض المقاييس الجوية . فلو قمنا بعمل توزيع تكراري لبضع مئات من قراءات ضغط الهواء في مدينة معينة فان التوزيع الناتج يكون ملتوياً التواء سالباً .



المنحنيات حادة الالتواء :

هناك نوعان من هذه المنحنيات ، منحنيات تتزايد التكرارات فيها كلما كبرت قيمة المتغير حتى تصل التكرارات اقصاها عند نقطة معينة من المتغير ولا تعود الى الانخفاض بعد ذلك كما هو واضح من الشكل الآتي :



ويسمى هذا النوع من المنحنيات منحنى رائي للشبه بينه وبين حرف الراء العربية ، وكذلك اطلق عليه في الانجليزية (J. Shaped Curve) للشبه بينه وبين حرف J .

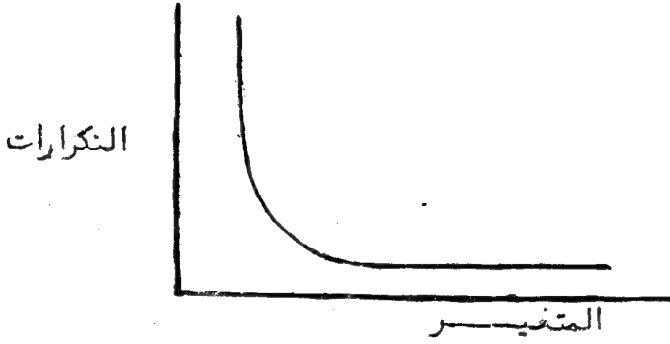
أما إذا كانت التكرارات كبيرة نوعاً ما عند مبدأ هذا المنحنى ثم تناقصت بعد ذلك فنطلق على هذا النوع من المنحنيات اسم منحنى لامبي تشبيهاً له بحرف اللام العربية ويلاحظ ان الكتاب الانجليزي لا يفرقون بين النوعين اللامي والرائي فيستخدمون اسماً واحداً لهما .

وهذا النوع من التوزيع التكراري نادر الحدوث في المسائل الاجتماعية . ولقد وجد أن التوزيع التكراري للمصابين في حوادث الطرق بحسب أعمارهم في إنجلترا من هذا النوع إذ ان عدد المصابين يتزايد كلما كبرت الأعمار حتى يصل العدد اقصاه عند الكهول وذلك طبعاً لضعف حواسهم الطبيعية ويقظتهم عند هذا العمر المتأخر . أما في مصر فقد حاولنا الوقوف على طبيعة شكل المنحنى لتلك الظاهرة نفسها فوجدنا ان المنحنى التكراري شبيه بمثله في

انجلترا مع اختلاف بسيط هو تزايد التكرارات نوعاً ما عند العمر ١٠ الى ١٥ سنة ، وهذا النوع من المنحنيات يطلق عليه اسم منحني لامي .

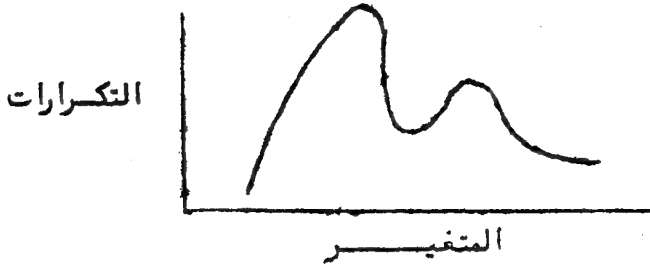
أما النوع الثاني من التوزيعات حاد الالتواء فهي ما كانت أكبر تكراراتها تقابل أصغر قيم المتغير ثم تبدأ التكرارات بعد ذلك في الانخفاض كلما زادت قيم المتغير حتى تصل التكرارات اقلها عند أكبر القيم كما هو مبين بالشكل التالي. ويطلق على هذا النوع من المنحنيات اسم رائي مقلوب Reversed J, shaped حيث انه عكس المنحنى الرائي السابق وصفه . أما إذا بدأت التكرارات تتزايد بعد وصولها الى نهايتها الدنيا فيطلق على هذا النوع اسم « منحني لامي مقلوب » . وهذا النوع من المنحنيات كثير الحدوث في المسائل الاقتصادية الخاصة بتوزيع الثروة ، مثلاً توزيع الملاك الزراعين بحسب مساحة ما يملكونه من أراض من أحسن الأمثلة لهذا النوع . ومن أمثلة هذا التوزيع أيضاً توزيع الشركات المساهمة حسب رؤوس أموالها وكذلك التوزيع التكراري لحالات الطلاق التي تحدث سنوياً في مصر موزعة حسب طول الحياة الزوجية التي سبقت الطلاق . ويكاد يكون التوزيع التكراري للسكان في مصر في الوقت الحاضر بحسب السن من هذا النوع ايضاً والشذوذ الوحيد في المثال الأخير هو أن عدد من هم دون الواحدة أقل نسبياً من عدد الفئة التي تليها مباشرة مما يجعل للمنحنى التكراري تقوس ضئيل إلى أسفل عند بدايته . ونلاحظ ان كثيراً ما نشاهد تلك الظاهرة (ظاهرة التقوس) في هذا النوع من المنحنيات كما نشاهد التقوس أيضاً في نهاية المنحنيات الرائية وكثيراً ما يعمد الاحصائيون إلى ضم الفئات الاولى من التوزيع إلى بعضها للتخلص من هذا التقوس .

ويسمى هذا المنحنى احياناً بمنحنى « باريتو » نسبة إلى وليفريد باريتو الذي وجه انتباه الاقتصاديين إلى هذا النوع في الدراسات الخاصة بتوزيع الثروة .



المنحنيات المتعددة القمم (Multi — Modal curves) :

وهي التي تتركز فيها القيم أي تعلو التكرارات عما حولها عند أكثر من نقطة واحدة كما هو واضح من الشكل الآتي :

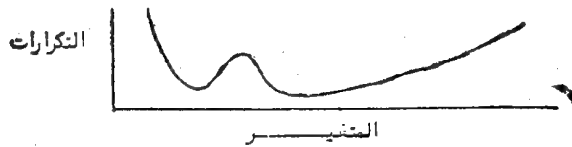


فإذا وجد توزيع تكراري من هذا النوع فإن ذلك دليل قاطع على عدم تجانس افراد المجموعة المكونة له . فلو فرض أن أخذنا مجموعة كبيرة من الأشخاص الذين ينتمون إلى أجناس مختلفة لم قسمنا أطوالهم وكونا توزيعاً تكرارياً من هذه الأطوال فإنه لا شك فيه ان التوزيع الناتج وبالتالي المنحني التكراري الذي يمثله يكون متعدد القمم حيث يتركز الطول لكل مجموعة عند نقطة معينة هي طول الغالبية العظمى لافراد كل جنس على حده . فلو فرض ان اقتصرنا في التوزيع على أطوال جنس واحد فإن التوزيع التكراري الناتج يفقد خاصية تعدد القمم ويصبح توزيعاً وحيد القمة ، فتعدد

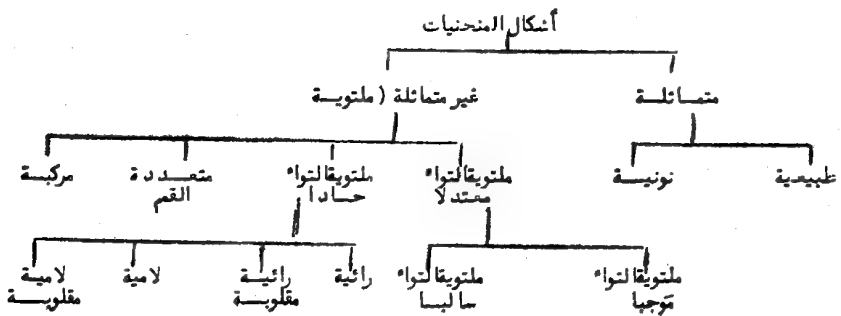
القمم في التوزيعات التكرارية اذن دليل مادي لا يخطيء على عدم تجانس المجموعة المكونة له . وكثيراً ما تستخدم هذه الحقيقة الهامة في البحوث العلمية للتحقق من تجانس المواد المختبرة أو عدم تجانسها .

فلو طلب منك ان تختبر نقاوة شحنة معينة من القمح مثلاً فيكفي أن تأخذ كمية محدودة من هذا القمح كعينة عشوائية ممثلة له وتكون توزيعاً تكرارياً لأوزان حبات القمح في تلك العينة ، فإذا كان المنحنى الذي يمثل أوزان هذه المجموعة منحني متعدد القمم كان ذلك دليلاً قاطعاً على أن القمح الذي تختبره مخلوط بنوع آخر من القمح وعدد القمم ينم على عدد الاصناف المخلوطة . أما إذا كان القمح نقياً ومن صنف واحد فمن المحتمل أن يكون التوزيع الناتج متاثلاً والمنحنى ذا قمة واحدة فقط .

وثمة نوع آخر من المنحنيات الملتوية وهو ما يطلق عليه المنحنيات المركبة (Complex Curves) وهو ما تكونت أجزاؤه من عدة أنواع مختلفة من المنحنيات التي ذكرناها سابقاً . فلو أخذنا التوزيع التكراري للمتوفين في انكلترا حسب أعمارهم عند الوفاة لوجدنا التوزيع كما هو مبين في الشكل الآتي :



ويمكن تلخيص التقسيم السابق في الشكل الآتي :



تمثيل التوزيعات التكرارية ذات الفئات الغير متساوية :

قلنا عند الكلام على اختيار فئات التوزيعات التكرارية بوجوب أخذ فئات متساوية. المدى إذ أن في ذلك تسهيل لمقارنة تكرارات التوزيع وتبسيط التمثيل البياني لها. ولكننا ذكرنا في معرض الكلام عن هذا الموضوع أن هناك اعتبارات عملية لا تسمح باتباع تلك النصيحة حيث أن طبيعة التوزيع تتطلب أحياناً التفصيل في بعض أجزائه والاختصار في البعض الآخر، الأمر الذي يقتضي جعل الفئات غير متساوية .

يترتب على ذلك أن هناك بعض القواعد التي يجب أن نتبعها عند مقارنة التوزيعات التكرارية ذات الفئات الغير متساوية وتمثيلها بيانياً . إذ أنه لا بد من تعديل تكرارات الفئات المختلفة تبعاً للمدى كل منها إذ لا يجوز بحال من الأحوال مقارنة تكرارات فئتين أو أكثر مداها مختلف دون تعديل .

والمثل الآتي يوضح ضرورة تعديل تكرارات الفئات الغير متساوية قبل مقارنتها .

لو قيل مثلاً ان عدد العمال الذين يتقاضون أجراً يومياً قدره يتراوح بين ٢٠ قرشاً و ٣٠ قرشاً هو ٥٠ عاملاً وان عدد ما يتقاضون أجراً بين ٣٠ ، ٦٠ قرشاً هو ١٢٠ عاملاً وسئلت بعد ذلك أي الفئتين أكبر عدداً - عمال الأجور الصغيرة أو عمال الأجور الكبيرة - فماذا يكون جوابك ؟ هل يصح أن تقول بما أن عمال الأجر العالي ١٢٠ عاملاً أكثر من عمال الأجر المنخفض ٥٠ عاملاً فان أصحاب الأجور المنخفضة أقل عدداً من أصحاب الأجور المرتفعة . الواقع أن عدد أصحاب الأجر المنخفض قليلون لا لسبب إلا لأن مدى فئتهم صغير (مدى ١٠ قروش فقط ، أي من ٢٠ إلى ٣٠ قرشاً) في حين أن مدى الأجور العالية هو ٣٠ قرش (أي من ٣٠ إلى ٦٠ قرشاً) . فلو فرض ان توزيع أصحاب الأجور العالية في فئتهم منتظم فمن الواضح ان عدد من يتقاضى أجراً من ٣٠ إلى ٤٠ قرشاً يساوي عدد من يتقاضى أجراً من ٤٠ إلى ٥٠ قرشاً يساوي عدد من يتقاضى أجراً من ٥٠ إلى ٦٠ قرشاً ، والعدد في كل

حالة هو ٤٠ عامل ومن ذلك نستنتج أن أصحاب الأجور المنخفضة في هذا المصنع أكثر عدداً من أصحاب الأجور العالية وهو عكس ما وصلنا اليه بمجرد مقارنة التكرارات الأصلية دون الاهتمام بمدى فئاتهم .
يستنتج من هذا انه لا بد من تعديل تكرارات الفئات الغير متساوية قبل مقارنة تلك التكرارات. وذلك بأن نأخذ طول فئة معينة كوحدة ثم نضرب تكرارات كل فئة في نسبة مدى الفئة المتخذة وحدة الى مدى تلك الفئة . فلو أخذنا مدى الفئة ١٠ في المثال السابق كوحدة ثم ضربنا تكرارات الفئة الثانية في النسبة بين مدى الوحدة ومدى تلك الفئة نحصل بذلك على التكرار المعدل .

$$\text{التكرار الأصلي} \times \frac{\text{طول الفئة المتخذة كوحدة}}{\text{طول الفئة المراد تعديلها}} = \text{التكرار المعدل}$$

$$120 \times \frac{10}{30} = 40$$

وإظهاراً لأهمية تلك القاعدة وتطبيقاً لها نعود إلى كشف البيانات الأصلية الخاص بأجور العمال ونكوّن منه توزيعاً تكرارياً تكون فئاته غير متساوية المدى ثم نمثل التوزيع بطريقتين الأولى خاطئة أن دون تعديل في التكرارات والثانية صحيحة وهي التي تم فيها هذا التعديل ، وسوف نرى اننا سنحصل على توزيع يختلف قانون تغيره اختلافاً تاماً عن قانون تغير التوزيع كما نعرفه . وسيرى القارئ أن تعديل التكرارات وفقاً للقاعدة التي ذكرناها سيعيد قانون التغير الى حقيقته .

فئات الأجور

٤ - ٦ - ٨ - ١٠ - ١٢ - ١٤ - ١٦ - ١٨ - ٢٠ - ٢٤ - ٣٠ - ٥٢

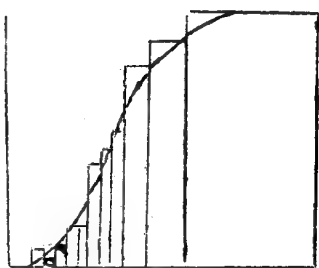
(عدد العمال) التكرار

٤ ٢ ٥ ٩ ٩ ٢٢ ٢٥ ٢٩ ٤٣ ٤٨ ٥٤

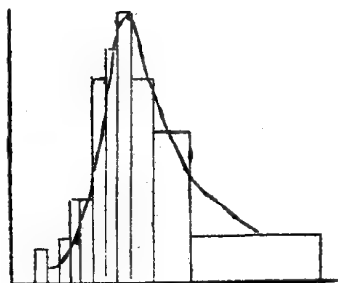
يذكر القارىء أن فائدة عمل توزيع تكراري تنحصر في ثلاث نقاط :
 الأولى تحديد مدى التغير ، والثانية الكشف عن وجود التركيز في القيم
 والقيمة التي يحدث عندها التركيز ، والثالثة الوقوف على الاستمرار في التغير
 بين حدى المدى .

وفضلاً عما تقدم فإن الرسم البياني على اختلاف أنواعه سواء كان أشرطة
 أو مدرجات أو مضلعات أو منحنيات يعطى للعين صورة مجسمة لقانون
 التغير في التوزيع موضوع الدراسة . ونعني بقانون التغير طريقته ، أي العلاقة
 بين التكرارات والتغير .

فإذا أخذنا توزيع الأجور السابق مثلاً نرى أنه كلما زاد الأجر زاد معه
 تكراره في بادىء الأمر حتى يصل التكرار الى أقصى حدوده عند حوالي ١٨
 ليرة ثم يبدأ التكرار يتناقص كلما زاد الأجر ولكن بسرعة أقل بكثير من
 السرعة التي ارتفع بها حتى وصل الى قمته . ومعنى ذلك أن الأجر في هذا
 المصنع « يميل » نحو الأجور الصغيرة ، إذ أن معظم التكرارات الكبيرة
 تلتف حول القيم الصغيرة للأجور .



- ١ -



- ٢ -

واضح من الرسم الاول (١) ان شكل المنحنى رائي وهذا يختلف
 اختلافاً بيناً عن الشكل الذي سبق أن حصلنا عليه للتوزيع نفسه عند
 تقسيمه الى فئات متساوية .

وغني عن القول ان التغيير في اختيار مدى الفئات لا يجب أن يؤثر على قانون تغير المجموعة او شكل منحناها . فذلك ثابت مهما اختلفت الفئات ، والسبب الحقيقي للنتيجة الخاطئة التي وصلنا اليها هو عدم تعديل التكرارات وفقاً للقاعدة التي شرحناها .

أما الجدول الآتي فيعطى التوزيع وتكراراته المعدلة . فاذا قمنا بتعديل التكرارات كما هو مبين في الجدول ورسمنا المدرج والمنحنى التكراري للتوزيع كما هو واضح بالشكل رقم (٢) فانتا نحصل على منحنى ملئو التواء موجباً وهو الشكل الحقيقي للتوزيع كما نعرفه .

فئات الأجر	التكرارات الأصلية (عدد العمال)	أطوال الفئات	معامل التصحيح طول الوحدة طول الفئة	التكرارات المعدلة التكرار الاصلي × معامل التصحيح
٤ -	٤	٢	١	٤
٦ -	٢	٢	١	٢
٨ -	٥	٢	١	٥
١٠ -	٩	٢	١	٩
١٢ -	٩	٢	١	٩
١٤ -	٢٢	٢	١	٢٢
١٦ -	٢٥	٢	١	٢٥
١٨ -	٢٩	٢	١	٢٩
٢٠ -	٤٣	٤	$\frac{٢}{٤}$	٢١,٥
٢٤ -	٤٨	٦	$\frac{٢}{٦}$	١٦
٣٠ - ٥٢	٥٤	٢٢	$\frac{٢}{٢٢}$	٤,٩
المجموع	٢٥٠	—	—	—

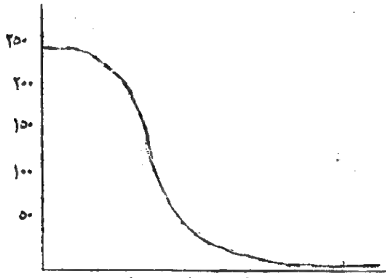
المنحنيات التكرارية المتجمعة :

ذكرنا سابقاً اننا نستفيد أحياناً بتجميع التكرارات اما صعوداً أو هبوطاً فنحصل على التوزيع التكراري المتجمع الصاعد أو الهابط . وتظهر

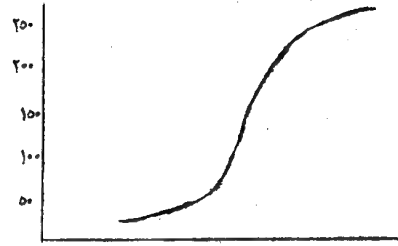
فائدة هذا النوع من التوزيعات بشكل واضح اذا مثلناها بيانياً كما فعلنا في التوزيعات التكرارية البسيطة .

وطريقة توضيح هذه التوزيعات تتلخص في تمثيل الحدود العليا للفئات على المحور الأفقي (المحور السيني) والتكرارات المتجمعة الصاعدة على المحور الرأسي (المحور الصادي) ثم نضع نقطة تمثل العلاقة بين حدود الفئات العليا وتكراراتها ثم نوصل النقط الناتجة بخط ممد فنحصل على الشكل الآتي .

ويطلق على المنحنى في هذه الحالة اسم المنحنى التكراري المتجمع الصاعد ويلاحظ بنوع خاص أن تكرارات كل فئة لا تقابل مراكز الفئات كما كان الحال في الموضع التكراري بل انها تقابل الحدود العليا للفئات . وبالمثل يمكن تمثيل المنحنى المنحنى المتجمع الهابط بيانياً بتخصيص المحور الأفقي للحدود الدنيا للفئات والمحور الرأسي للتكرارات المتجمعة الهابطة ثم تحديد النقط الدالة على العلاقة بين الحدود الدنيا والتكرارات من واقع الجدول ثم توصيل تلك النقط بخط ممد كما هو مبين بالشكل ويطلق على هذا المنحنى اسم المنحنى التكراري المتجمع الهابط، ويجب الإشارة هنا كذلك الى أن تكرارات كل فئة تقابل الحدود الدنيا للفئات وليست مراكزها كما كان الحال في الرسوم البيانية التكرارية البسيطة .

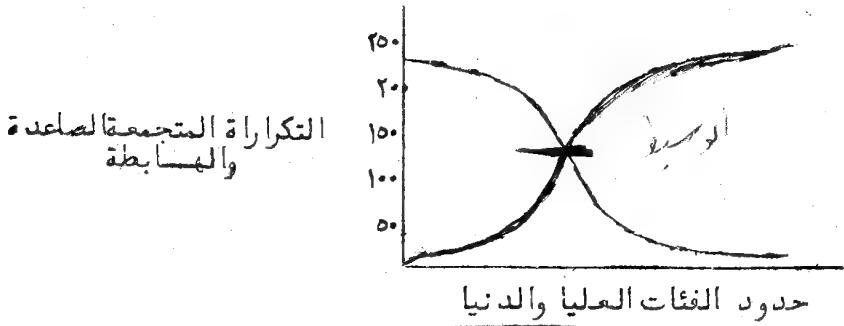


الحدود الدنيا للفئات
المنحنى الهابط -



الحدود العليا للفئات
المنحنى الصاعد -

ولا يفوتنا أن نذكر أنه من الممكن تمثيل المنحنيان المتجمعان الصاعد والهابط في شكل واحد كما هو مبين بالرسم التالي . ويلاحظ أن نقطة تقاطع هذين المنحنيين ذات أهمية خاصة إذ يقطع العمود النازل منها المحور الأفقي في نقطة تمثل أحد متوسطات الأجر، إذ أنها النقطة الوحيدة التي يتساوى عندها عدد الذين يتقاضون أجراً أقل منها وعدد الذين يتقاضون أجراً أكبر منها ، وسنطلق على هذا النوع من المتوسطات اسم « الوسيط » (Median) كما سنشرح ذلك عند الكلام على موضوع المتوسطات .



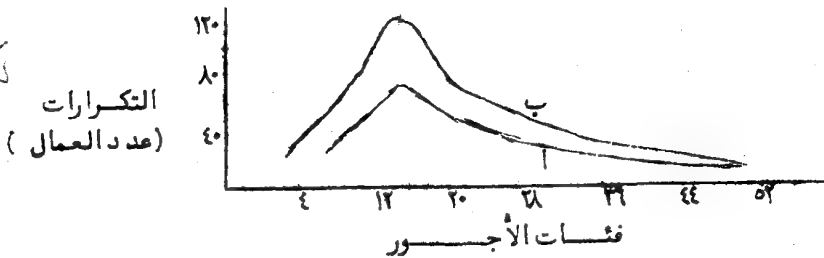
وفائدة هذه المنحنيات هي نفس فائدة التوزيعات المتجمعة فيمكن منها معرفة عدد العمال الذين يتقاضون أقل من أجر معين أو أكثر من أجر معين ، كما يمكن معرفة الأجر الذي يتقاضى دونه أو أكثر منه نصف العمال أو ربعهم أو أي عدد معين من العمال .

مقارنة التوزيعات التكرارية :

لو طلب منا مقارنة التوزيع التكراري لأجور عمال مصنعين كالآتي :

عمال مصنع ب	عمال مصنع أ	فئات الأجور
٧	٦	— ٤
٢٥	١٤	— ٨
٦٨	٣١	— ١٢
١١٦	٥٤	— ١٦
٩٦	٤٣	— ٢٠
٨٤	٣٦	— ٢٤
٦٨	٢٧	— ٢٨
٥٦	١٩	— ٣٢
٤٠	١٠	— ٣٦
٢٤	٦	— ٤٠
١٣	٣	— ٣٤
٣	١	٥٢ — ٤٨
٦٠٠	٢٥٠	

فخبر الوسائل المتبعة لمقارنة مثل هذين التوزيعين هو تمثيلها في شكل بياني . ويجب استخدام المنحنى التكراري دون المدرج التكراري إذ يمكن رسم المنحنيين على نفس الشكل .



واضح من الشكل السابق أن مدى التغير في الأجور واحد تقريباً في المصنعين وان التغير مستمر في الحالتين بين حدى المدى وان نقطة التركيز تكاد تكون واحدة حيث انها تتراوح بين ١٤ ، ١٨ ليرة في المصنعين . ولكن بالنظر إلى الشكل لا يمكننا مقارنة نسبة عدد العمال الذين يتقاضون أجوراً معينة إذ أن الشكل يحملنا على الاعتقاد لأول وهلة ان نسبة العمال الذين يتقاضون فئات الأجور المتوسطة أكبر بين عمال مصنع ب منها بين عمال مصنع أ ، ولكن الواقع غير ذلك إذ أن نسبة العمال الذين يتقاضون الأجور الصغيرة والمتوسطة بين عمال مصنع أ أكبر منها بين عمال مصنع ب وعلى العكس نسبة العمال الذين يتقاضون أجوراً عالية في مصنع أ أقل منها في مصنع ب . ولا يمكننا الوقوف على تلك الحقائق إلا بعد تحويل أرقام التوزيعين السابقين إلى نسب مئوية حيث ان مجموع عدد العمال في العيقتين غير متساو . والجدول التالي يعطي التوزيعين بتكرارتهما العادية والمئوية بالنسبة للمجموع في كل مصنع .

فئات الأجور	عمال مصنع أ	النسب المئوية لهم	عمال مصنع ب	النسب المئوية لهم
٤ -	٦	٢,٤	٧	١,١٦
٨ -	١٤	٥,٦	٢٥	٤,١٦
١٢ -	٣١	١٢,٤	٦٨	١١,٣
١٦ -	٥٤	٢١,٦	١١٦	١٩,٣
٢٠ -	٤٣	١٧,٢	٩٦	١٦,٠
٢٤ -	٣٦	١٤,٤	٨٤	١٤,٠
٢٨ -	٣٧	١٥,٨	٦٨	١١,٣
٣٢ -	١٩	٧,٦	٥٦	٩,٣
٣٦ -	١٠	٤,٠	٤٠	٦,٦
٤٠ -	٦	٢,٤	٢٤	٤,٠
٤٤ -	٣	١,٢	١٣	٢,١٦
٤٨ - ٥٢	١	٠,٤	٣	٠,٤٩
	٢٥٠	١٠٠	٦٠٠	١٠٠

١٨ × ١/٥

١١ × ١/٥
٩ × ١/٥

١ × ١/٥

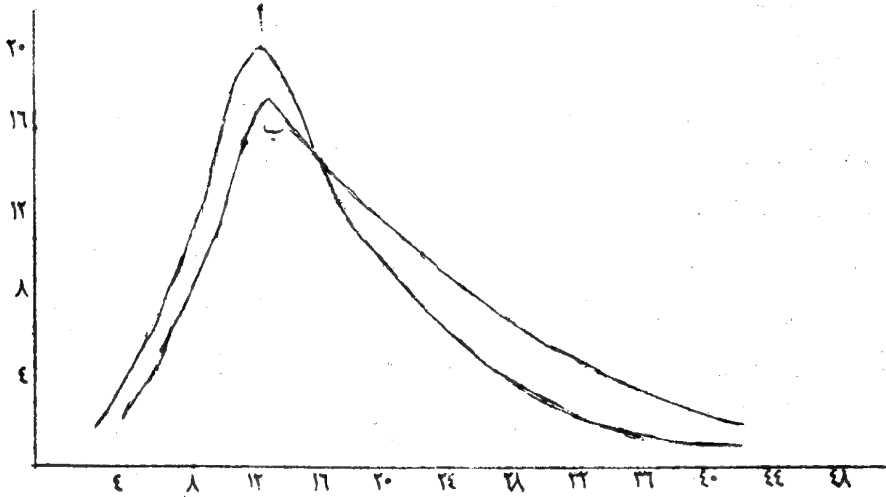
١ × ١/٥

١٨ × ١/٥

١٨ × ١/٥

١٨ × ١/٥

فاذا أوضحنا التوزيع التكراري للنسب المئوية في كل من المصنعين نحصل على الرسم الآتي :



وبلاحظ القارئ أن جميع فئات التوزيعين التكراريين السابقين متساوية، فاذا فرض أن كان لدينا توزيعين فئاتها غير متساوية المدى فانه من الواجب تعديل التكرارات الأصلية وبالتالي النسب المئوية لتلك التكرارات تبعاً لمدى كل فئة بالطريقة المعتادة قبل عمل الرسم البياني .

كذلك يلاحظ القارئ أن المقارنة تمت بين توزيعين طبيعة مميز المتغيرين في كل منها واحدة (الليرة مثلاً) . ولكن لا يمكن عمل مقارنة من هذا النوع بين توزيعين يختلف مميز المتغير فيها .

وقد يطلب منا في بعض الأحيان مقارنة توزيعين تكراريين متجمعين ، وفي هذه الحالة لا يجب الاعتماد على المنحنيين المتجمعين بل يلزم تحويل التكرارات المتجمعة لكل من التوزيعين الى نسب مئوية ورسم المنحنيين من واقع هذه النسب . ويلاحظ عند مقارنة التوزيعات التكرارية المتجمعة انه لا تأثير لاختلاف مدى الفئات على التكرارات المقارنة بعكس الحال في مقارنة التوزيعات التكرارية البسيطة . أما فيما يختص بمقارنة التوزيعات التكرارية المتجمعة ذات المتغيرات المختلفة فانه يستحيل علينا مقارنة منحنياتها المتجمعة أسوة بمنحنياتها العادية .

منحنى لورنز :

يمكننا الاستفادة من عمل التوزيع التكراري لأي متغير يجانب ما ذكرنا في الوقوف على درجة « التفاوت » بين أفراد المجموعة وذلك بتمثيل التوزيع بمنحنى جديد يطلق عليه اسم منحنى « لورنز » (Lorenz Curve) اشارة باسم مبتكره .

لنفرض اننا نريد عمل منحنى لورنز لتوزيع الأجور المعروف فينحصر العمل في الخطوات التالية :

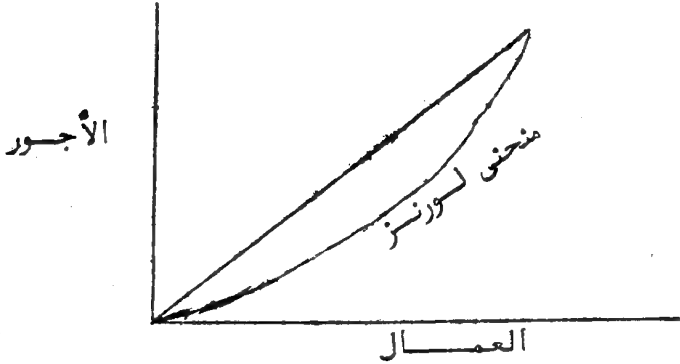
١ - ايجاد الأجور التي يتقاضاها أفراد كل فئة على حدة بضرب عدد عمال الفئة في مركزها .

٢ - عمل توزيعين متجمعين أحدهما للتكرارات الأصلية (أي العمال) والآخر للأجور التي يتقاضا هؤلاء العمال (حاصل ضرب تكرار كل فئة \times مركزها) .

٣ - حساب النسب المئوية لكل فئة من فئات التوزيعين ثم رصد هذه النسب على المحورين الرأسى والأفقى فيكون الخط الذي يمثل العلاقة بينهما هو منحنى لورنز المطلوب .

فئات الأجور	التكرارات	المراكز	التكرارات \times المراكز	التكرارات متجمعة	الأجور متجمعة	تكرارات متئوية متجمعة	أجور متئوية متجمعة
٤ -	٦	٦	٣٦	٦	٣٦	٢٠٤	٠٠٦
٨ -	١٠	١٠	١٤٠	٢٠	١٧٦	٨٠٠	٣٠١
١٢ -	١٤	١٤	٤٣٤	٥١	٦١٠	٢٠٠٤	١٠٠٦
١٦ -	١٨	١٨	٦٧٢	١٠٥	١٥٨٢	٤٢٠٠	٢٧٠٦
٢٠ -	٢٢	٢٢	٩٤٦	١٤٨	٢٥٢٨	٥٩٠٢	٤٤٠٠
٢٤ -	٢٦	٢٦	٩٣٦	١٨٤	٣٤٦٤	٧٣٠٦	٦٠٠٣
٢٨ -	٣٠	٣٠	٨١٠	٢١١	٤٢٧٤	٨٤٠٤	٧٤٠٥
٣٢ -	٣٤	٣٤	٦٤٦	٢٣٠	٤٩٢٠	٩٢٠٠	٨٥٠٧
٣٦ -	٣٨	٣٨	٣٨٠	٢٤٠	٥٣٠٠	٩٦٠٠	٩٢٠٣
٤٠ -	٤٢	٤٢	٢٥٢	٢٤٦	٥٥٥٢	٩٨٠٤	٩٦٠٧
٤٤ -	٤٦	٤٦	١٣٨	٢٤٩	٥٦٩٠	٩٩٠٠	٩٩٠١
٤٨ - ٥٢	٥٠	٥٠	٥٠	٢٥٠	٥٧٤٠	١٠٠٠٠	١٠٠٠٠
المجموع	٢٥٠	-	٥٧٤٠	-	-	-	-

لو لم يكن هناك تفاوتاً بين نسبة العمال ونسبة الأجور التي يتقاضونها بمعنى أن ٢٠٪ من العمال يتقاضون ٢٠٪ من الأجور وأن ٤٠٪ من العمال يتقاضون ٤٠٪ من الأجور وهكذا، لانطبق منحنى لورنز على قطر المربع المرسوم فيه كما



هو واضح في الرسم ، ولكن نظراً لأن الحقيقة غير ذلك إذ أن ٢٠٤ من العمال يتقاضون ١٠٦٪ من مجموع الأجور وأن ٤٢٪ من العمال يتقاضون ٢٧٦٪ من الأجور فإن المنحنى تقوس الى أدنى بعض الشيء وبدل مقدار التقوس على مدى التفاوت بين التوزيعين - العمال وأجورهم - ويمكن تظليل المساحة بين الخط المستقيم والمنحنى لإظهار مدى التفاوت في التوزيعين فكلما زادت هذه المساحة المظلة كلما دل ذلك على تفاوت أكبر .

ويستخدم هذا المنحنى بكثرة في التوزيعات التكرارية التي يهتم الباحث الوقوف على درجة التفاوت فيها . فبينما في بعض الحالات مقارنة عدد المصانع وكمية ما تنتجها لإظهار تركيز الانتاج في وحدات قليلة أو انعدام هذا التركيز وكذلك توزيعات الدخل بين أفراد المجتمع لإظهار تركيز الثروة في يد أفراد قلائل أو انعدام هذا التركيز .

وكذلك يمكن عمل منحنين من هذا النوع ورسمها في شكل واحد لمقارنة مدى التفاوت في التوزيعين اللذين نقوم ببحثهما . (قارن التفاوت في الأجور في المصنعين أ ، ب) . ونلاحظ أنه إذا تعادل التفاوت في التوزيعين فيستحسن رسم المنحنين على جانبي خط التماثل وذلك باستخدام المحور الأفقي

للمعال في المصنع أ والمحور الرأسي لأجورهم ، ثم استخدام المحور الأفقي للأجور والرأس للمعال في المصنع ب ، وذلك لامكان عمل المقارنة بينهما ، أما إذا كانت نسبة التفاوت تختلف كثيراً بين التوزيعين يمكن رسم المنحنيين على جانب واحد من خط التماثل حيث تسهل المقارنة في هذه الحالة .

التصوير الجبري للخطوط البيانية :

سبق أن ذكرنا ان الخط الذي يتوسط النقط التي نرصدها في الرسم البياني يسمى خط الاتجاه العام للظاهرة موضوع البحث ، حيث انه يدل على اتجاه التغير في قيم هذه الظاهرة في الفترات الزمنية المختلفة . كذلك إذا كان الرسم يوضح ظاهرتين سوياً فان خط الاتجاه العام يكون هو الخط الذي يمثل اتجاه التغير في احدي الظاهرتين تبعاً للتغير في الظاهرة الاخرى . ومن الواضح أن نوع التغير في الظاهرة سواء كان تغيراً زمنياً أو تغيراً ناتجاً عن تغير ظاهرة اخرى لا بد أن يؤثر على شكل الخط الذي يمثل الاتجاه العام .

وخط الاتجاه العام يمكن أن يكون خطاً مستقيماً اما صاعداً نحو اليمين أو هابطاً نحو اليمين ويدل على أن نوع التغير في الظاهرة موضوع البحث ثابتاً دائماً فهي اما في ازدياد مستمر إذا كان صاعداً نحو اليمين أو نقص مستمر اذا كان هابطاً نحو اليمين . ان هذا الشكل المنتظم لخط الاتجاه العام جعل الرياضيين يحاولون وضع معادلة تدل عليه وقد توصلوا الى ان معادلة من الدرجة الاولى تعبر عن اتجاهه ، وتتخذ هذه المعادلة الصيغة الآتية : $v = m + s$ حيث m تدل على ميل الخط وتكون موجبة إذا كان الخط صاعداً نحو اليمين وسالبة إذا كان الاتجاه العام هابطاً نحو اليمين ، أما b فقيمة ثابتة تدل على الجزء الذي يقطعه خط الاتجاه العام من المحور الرأسي .

وعندما يكون الخط معبراً عن الاتجاه العام للظاهرة في فترات زمنية مختلفة تكون s في المعادلة رمزاً لهذه الفترات ، أما إذا كان الخط معبراً عن

الاتجاه العام للعلاقة بين ظاهرتين تكون س ، ص رموزا لهما وفي هذه الحالة يمكن أن يوجد خطان للاتجاه العام ، احدهما يصور اتجاه الظاهرة ص في علاقتها مع الظاهرة الاخرى س ، ويظهر في المعادلة $ص = م س + ب$ ، والثاني يصور اتجاه الظاهرة س في علاقتها مع الظاهرة ص ، ويظهر في المعادلة $س = م ص + ب$ وفي هذه المعادلة تكون م هي كذلك رمزاً لميل خط اتجاه س بالنسبة للظاهرة ص ، ب رمزا للجزء المقطوع من المحور س بواسطة خط الاتجاه العام .

وقد سبق أن ذكرنا أن خط الاتجاه العام يجب أن يتوسط النقط التي تمثل الاتجاه العام للظاهرة في تغيرها الزمني أو الاتجاه العام للعلاقة بين ظاهرتين ، ومن الواضح أنه من الصعب رسم مثل هذا الخط بحيث يتوفر فيه شرط التوسط ، الا انه على أساس معادلة الاتجاه العام يمكن تحديد النقط التي تقع عليه وبذلك يمكن تحديد اتجاهه الوسطي بين النقط الواقعية بدقة تامة . والمشكلة هي في كيفية إيجاد هذه المعادلة ويستخدم لذلك ما نسميه بطريقة المربعات الصغرى حيث أننا بهذه الطريقة نوجد معادلة أفضل خط يمثل الاتجاه العام وهو الخط الذي يتوفر فيه الشرطان :

١ - ان مجموع انحرافات النقط الواقعية عن هذا الخط = صفر .

٣ - ان مجموع مربع انحرافات النقط الواقعية عند هذا الخط أصغر من مجموع مربعات انحرافات النقط الواقعية عن أي خط آخر .

وسوف نعود الى مناقشة هذه الطريقة في موضوع الانحدار .

كذلك يمكن أن يكون الاتجاه العام متخذاً شكل منحني من الدرجة الثانية وتصوره المعادلة $ص = أ س^٢ + ب س + ج$ حيث أ ، ب ، ج ثوابت معينة . وتمثل هذه المعادلة ما نسميه بالقطع المكافئ ، وهو اما محدباً الى أعلى أو مقعراً الى أعلى ، فاذا كانت قيمة أ سالبة يكون القطع المكافئ

معدباً الى أعلى ، اما اذا كانت موجبة يكون القطع مقعراً الى أعلى . وسواء كانت موجبة أو سالبة فان النقطة التي يصل القطع المكافئ الى قيمته الكبرى أو الصغرى عندها هي النقط التي تتحدد على أساس المعادلين :

$$س = -\frac{ب}{١٢} ، ص = \frac{٢١٤ - ب}{١٤} . هذه النقطة تسمى رأس القطع$$

المكافئ. ومن الواضح أن الشكل العام لمعادلة الدرجة الثانية يمكن أن يختلف تبعاً لما اذا كان القطع المكافئ يقطع محور س مرتين واما أن يقطعه مرة واحدة أي يمسه، واما أن لا يقطعه بتاتاً فيقع بتمامه فوق محور س أو تحته .

ومنحنى القطع المكافئ يجب كذلك أن يتوسط النقط التي تمثل الاتجاه العام ولهذا يمكن استخدام طريقة المربعات الصغرى في إيجاد معادلته .

كذلك يمكن أن يكون الاتجاه العام من الدرجة الثالثة فتمثله المعادلة .

$$ص = أ س^٣ + ب س^٢ + ج س + د$$

كذلك يمكن أن يكون الاتجاه العام من درجة أعلى .

ومن الواضح أن المنحنيات التكرارية يمكن التعبير عنها بمعادلات وستترك ذلك في الوقت الحاضر حتى ندرس الخصائص الرياضية للتوزيع التكراري .

القواعد العامة للرسم البياني :

الرسم البياني هو كما ذكرنا الوسيلة الثانية التي يلجأ اليها الاحصائي لتوضيح البيانات بعد جمعها وتبويبها حيث أن الأرقام المبوبة في جداول قد تكون صعبة الفهم والاستيعاب بالنسبة لكثير من الناس . لذلك يجب أن يكون الرسم البياني واضحاً من جميع النواحي حتى يمكن أن يؤدي وظيفته التوضيحية بالنسبة لمن يطلع عليه .

كذلك تساعد بعض أنواع من الرسومات البيانية في اجراء تحليلات احصائية معينة . وفي هذه الحالة يجب أن يتوفر في الرسم الدقة (ليست الدقة الهندسية الكاملة) حيث ان استنتاجنا من الرسم عندما لا يكون دقيقاً يمكن أن يكون خاطئاً مضللاً .

لذلك يكون الوضوح التام والدقة الممكنة هما الصفتان الأساسيتان في الرسوم البيانية الاحصائية . وحتى يمكن أن يتوفر في الرسم هاتان الصفتان يجب أن نلاحظ القواعد الآتية عند تصميمه وتنفيذه :

أولاً :

لكل نوع من البيانات الاحصائية المبوبة رسم خاص بها يكون قادراً على توضيحها و ابراز الفكرة الكامنة وراءها ، حيث اننا بأي رسم بياني نريد أن نقول لمن يطلع عليه شيئاً معيناً . هذا الشيء هو الفكرة الكامنة وراء البيانات التي نقوم بتوضيحها . قد يكون غرضنا من الرسم أن نبرز طريقة التغير في الظواهر موضوع البحث خلال فترة زمنية معينة ، وقد يكون غرضنا أن نظهر تقسيم بيان كلي إلى أجزائه المختلفة في عام واحد ، وقد يكون غرضنا أن نعرض أن نظهر هذا التقسيم في عدة أعوام ليتضح التغير في تركيب الظاهرة من عام الى آخر ، وقد يكون غرضنا أن نظهر المقارنة في طريقة التغير بين ظواهر مختلفة خلال فترة زمنية معينة ، وقد يكون غرضنا أن نظهر مقدار ما حدث من زيادة أو نقصان في الظاهرة أو الظواهر موضوع البحث ، وقد يكون غرضنا أن نظهر معدل ما حدث من زيادة أو نقصان في هذه الظواهر . الخ . لذلك تكون أول خطوة في تصميم الرسم البياني الاحصائي هي تحديد نوع الرسم الذي يتفق مع البيانات التي تريد توضيحها ومع الفكرة التي نريد ابرازها . ان الهدف من الرسم لا يمكن أن يتحقق تماماً الا باختيار النوع المناسب .

ثانياً :

ان الفكرة الأساسية من الرسم البياني الاحصائي هي التوضيح لذلك يجب أن يتوفر فيه السهولة والذوق السليم من جميع النواحي . فإذا كان الرسم يحتوي على خطوط بيانية مثلاً يجب أن تظهر ثقيلة واضحة بحيث تبرز للعين دون حاجة الى فحص وتدقيق (وذلك بعكس الرسوم البيانية الهندسية حيث تكون الخطوط فيها رقيقة خفيفة) . كذلك إذا احتوى الرسم على عدة خطوط بيانية يجب التميز بينها بالألوان أو باي طريقة اخرى وفي هذه الحالة يجب أن يصحب الرسم توضيح لمعاني هذه الألوان أو الطرق المختلفة التي استعملت . كذلك إذا أردنا أن نستعمل الألوان لا يجب أن تفعل ذلك الا إذا كان لدينا خبرة تامة باستعمالها . وبشكل عام كل ما يساعد في جعل الرسم مشوقاً يجب استعماله .

ثالثاً :

إن القيم التي نريد توضيحها بالرسومات البيانية الاحصائية قد لا يمكن أن نظهرها بحيث يتوفر فيها الدقة الحسابية التامة ولذلك نضطر أحياناً الى تقريب الأرقام حتى نظهرها ، إذ المقصود من أي رسم احصائي هو اعطاء فكرة توضيحية لمن يطلع عليه ، وليس اظهار القيم بدقة حسابية متناهية فإذا كان لدينا القيمة ٥٨٣٧٤٥٤٣ مثلاً نلاحظ أنه ليس من المعقول أن نحاول اظهار هذه القيمة بكاملها على الرسم البياني العادي وبذلك نضطر الى تقريبها الى أقرب مائة ألف حتى نستطيع أن نحدد موقع النقط التي تعبر عنها تبعاً لمقياس الرسم الذي نتخذه .

رابعاً :

يجب البدء بالصف على المحور الرأسي إلا إذا كان الرسم نصف لوغاريتمي ، وكذلك يجب أن يستمر مقياس الرسم منتظماً ، أما إذا تغير

مقياس الرسم يجب الاشارة الى ذلك في مكانه باستخدام علامة مميزة // .
وبشكل عام لا يجب أن نغير من مقياس الرسم على المحور الرأسي إذا كان
جزء من الرسم يدخل ضمن المقياس المختلف ، وبمعنى آخر إذا اضطررنا الى
ذلك تبعاً لطبيعة الأرقام يجب أن نلاحظ البدء بالصفر ثم نضع العلامة //
بحيث تكون أقل قيمة نريد توضيحها لا تقل بأي حال عن القيمة التالية التي
نضعها بعد هذه العلامة (//) . أما المحور الأفقي حيث انه غالباً يعبر عن
الزمن فيمكن البدء بأي سنة نريدها بحيث يتسع فراغ الرسم لجميع السنوات
المطلوب عرض البيانات الخاصة بها .

خامساً :

يجب أن تبرز النقط التي تحدد اتجاه الخط البياني بحيث تكون أكثر
وضوحاً من الخط نفسه .

سادساً :

يجب أن نتحاشى توضيح بيانات كثيرة ومتنوعة في رسم بياني واحد
لان ذلك يؤدي الى التعقيد وعدم الوضوح وبذلك يفقد الرسم ميزته .

سابعاً :

إذا كان الرسم خاصاً بفترة زمنية تبدأ بسنوات وتنتهي بأشهر السنة
الأخيرة يجب وضع فاصل في نهاية البيانات السنوية يكون موازياً للمحور
الرأسي وبذلك يكون الرسم في الوقع مكون من جزأين ، الجزء الأول
ويظهر البيانات السنوية والجزء الثاني ويظهر للبيانات الشهرية . وتستخدم هذه
الطريقة بشكل عام إذا اختلفت الوحدات الزمنية في السلسلة التي نريد توضيحها .

ثامناً :

يجب توضيح عنوان الرسم توضيحاً يظهر الموضوع والتاريخ والمكان

المتعلق بالرسم ، وبشكل عام يكون عنوان الرسم هو نفس عنوان الجدول الذي يحتوي على البيانات التي نؤمن بتوضيحها .

تاسعاً :

يجب الإشارة في أسفل الرسم الى مصدر البيانات التي يعتمد عليها . ويجب ذكر المصدر متضمناً اسم الكتاب واسم المؤلف ورقم الصفحة وتاريخ النشر ، وكذلك الحال إذا كان المصدر نشرة احصائية .

عاشراً :

يجب توضيح مبرز كل من المحورين - الأفقي والرأسي - بحيث يظهر على كل منها دلالة الخاصة به

حادي عشر :

يجب توضيح مقياس الرسم في مكان واضح من الشكل البياني :

تقارن

١ - تريد إحدى المؤسسات الاحصائية القيام بدراسة عن معيشة الطالـب الجامعي في بيروت . اكتب تقريراً توضح فيه الخطوات التي يجري العمل تبعاً لها ، مع تصميم الاستمارة التي ترى استخدامها لجمع المعلومات .

٢ - ناقش الأخطاء المختلفة التي تتعرض لها الدراسات الاحصائية وبين كيف يمكن التغلب عليها .

٣ - اكتب ما تعرفه عما يأتي :

المجتمع الاحصائي والوحدة الاحصائية - ، مجتمع المعاينة ووحدة المعاينة - دراسة احصائية وصفية بالعد الشامل - تضليل البيانات الاحصائية - الأساس

القانوني للدراسات الاحصائية - الرمم البياني كوسيلة لتوضيح البيانات الاحصائية .

٤ - في دراسة عن ميزانية الاسرة اشتملت استمارة البحث على سؤال عن قيمة الانفاق على الطعام وقد جاءت الاجابات كالآتي (بالدينار) .

٣٢	١٢	٤٥	٥٣	٤٤	١٨	١٣	٢٥	٨	٣٥	١٠
١٦	٢٦	٣٦	٤٣	١١	٥٢	٩	١٤	٤٠	٢٨	٢٧
٢٥	٢٢	١٠	٣٤	٣٦	١٥	١٤	٤١	٥٤	١٧	٥
٢٠	٧	٢٩	٢٨	٢٧	٢٠	٢٤	١٢	٢٣	٢١	٢٧
٤٢	٥١	٣٠	٢٥	٢٢	٢٧	٤١	٢١	٣١	٢٧	٢٣
١٧	٤١	٥١	٤٤	٢٠	٢٥	٢٦	٢١	١٥	٢٣	٣٧
٢٩	٢٦	٢٠	٢١	٢٣	٤٨	٢١	١٦	٢٢	٢٨	٢٥
٥٠	٤١	١٨	٣٠	٢١	٩	٢٦	٢٧	٢٧	٢٨	٢١
٢٦	٢٧	٥٠	٢٣	٢٥	٢٧	٣٠	٢٨	٢٩	٢٤	٢٢
										٢٩

المطلوب :

تبويب هذه البيانات في توزيع تكراري تبدأ فئاته بالقيمة ٥ ومدى الفئة فيه ٥ .

٥ - يجب قبل البدء في أي تعداد اتخاذ بعض القرارات الأساسية . ما هي هذه القرارات ؟ ولماذا يجب اتخاذها أولاً ؟ وضع اجابتك بأمثلة .

٦ - الآتي توزيع السكان تبعاً لفئات العمر في إحدى الدول

<u>فئات العمر</u>	<u>ذكور</u>	<u>إناث</u>
أقل من ٥	٦٦٧٧	٦٦١٦
٥ - ٩	٥٢٣٦	٥٢٢٨
١٠ - ١٤	٤٤٧٠	٣٥٧٧
١٥ - ١٩	٥١٧٧	٣٣٦١
٢٠ - ٢٤	١٠٣٠٨	٣٤٣٤
٢٥ - ٢٩	٩١٢٤	٣١٧٥
٣٠ - ٣٤	٦٣٨٠	٢٣٠٠
٣٥ - ٣٩	٤٣٠٠	١٩١١
٤٠ - ٤٤	٣٢٢٨	١٥١٨
٤٥ - ٤٩	٢٢٧٧	١١١٧
٥٠ - ٥٤	١٩٤١	١٢٤٣
٥٥ - ٥٩	٨٧٦	٦٢٨
٦٠ - ٦٤	٩٣٢	٧٩٤
٦٥ - ٦٩	٣٩٧	٣٠٩
٧٠ فأكثر	٨١٣	٨٨٣

ارسم هرم التركيب العمري وبين ملاحظاتك عليه .

٧ - جمعت البيانات الآتية عن ٥٠ مؤسسة تجارية . موقع المؤسسة ورأس مالها بالليرة والضرائب التي دفعتها في عام (بالليرة) .

بيروت	الشمال	الجنوب	بيروت
٥٠٠٠٠ - ٤	٢٥٠٠٠ - ٣	٣٠٠٠٠ - ٢	٢٠٠٠٠ - ١
١٢٠٠	٦٥٠	٧٠٠	٥٠٠
بيروت	الشمال	الجنوب	بيروت
٣٠٠٠٠ - ٨	١٥٠٠٠ - ٧	١٠٠٠٠ - ٦	١٥٠٠٠ - ٥
٧٥٠	٤٠٠	٣٥٠	٤٥٠
الشمال	بيروت	الجنوب	الشمال
٣٥٠٠٠ - ١٢	٧٠٠٠٠ - ١١	١٦٠٠٠ - ١٠	١٥٠٠٠ - ٩
٩٠٠	١٥٠٠	٥٠٠	٤٥٠
الجنوب	بيروت	الشمال	بيروت
٤٠٠٠٠ - ١٦	٤٥٠٠٠ - ١٥	٦٥٠٠٠ - ١٤	٨٥٠٠٠ - ١٣
٧٥٠	٨٠٠	١٢٠٠	١٧٠٠
بيروت	الجنوب	الشمال	بيروت
٥٥٠٠٠ - ٢٠	٦٥٠٠٠ - ١٩	٣٥٠٠٠ - ١٨	١٨٠٠٠ - ١٧
٨٥٠	٩٠٠	٨٠٠	٤٥٠
بيروت	الجنوب	الشمال	بيروت
١٩٠٠٠ - ٢٤	٧٥٠٠٠ - ٢٣	١٢٠٠٠ - ٢٢	٣٥٠٠٠ - ٢١
٦٠٠	١١٠٠	٣٠٠	٥٠٠
الجنوب	الشمال	الجنوب	الشمال
١٨٠٠٠ - ٢٨	٤٥٠٠٠ - ٢٧	٣٨٠٠٠ - ٢٦	٧٠٠٠٠ - ٢٥
٦٠٠	٩٠٠	٨٠٠	٩٠٠

بيروت	بيروت	الشمال	الشمال
٢٧٠٠٠ - ٢٩	٩٠٠٠٠ - ٣١	٥٥٠٠٠ - ٣٠	٦٠٠٠٠ - ٣٢
٧٥٠	١٦٠٠	٩٠٠	٩٥٠

الجنوب	الشمال	بيروت	بيروت
٤٨٠٠٠ - ٣٣	٤٢٠٠٠ - ٣٥	١٧٠٠٠ - ٣٤	١٤٠٠٠ - ٣٦
٧٥٠	٩٥٠	٦٥٠	٣٥٠

الشمال	الجنوب	بيروت	الشمال
١٨٠٠٠ - ٣٧	٧٣٠٠٠ - ٣٩	٣٥٠٠٠ - ٣٨	١٦٠٠٠ - ٤٠
٦٥٠	١١٠٠	٩٥٠	٥٥٠

بيروت	الجنوب	الشمال	بيروت
٣٣٠٠٠ - ٤١	٢٣٠٠٠ - ٤٣	١٩٠٠٠ - ٤٢	١٧٠٠٠ - ٤٤
٩٥٠	٧٥٠	٥٥٠	٦٠٠

الشمال	الجنوب	بيروت	بيروت
٣٨٠٠٠ - ٤٥	٤٣٠٠٠ - ٤٧	٤٧٠٠٠ - ٤٦	١٣٠٠٠ - ٤٨
٩٥٠	٩٥٠	١٠٠٠	٤٠٠

بيروت	الجنوب
٢٧٠٠٠ - ٤٩	١٨٠٠٠ -
٨٠٠	٥٠٠

المطلوب :

١ - تبويب المؤسسات التجارية ورؤوس اموالها والضرائب التي دفعتها
تبعاً للموقع وفئات رأس المال بحيث يظهر الجدول المجاميع الخاصة
بكل منطقة على حدة ومتوسطات رأس المال والضرائب .

الموقع - بيروت - الشمال - الجنوب

فئات رأس المال - ١٠٠٠٠ - ٢٠٠٠٠ - ٥٠٠٠٠ وأكثر

٢ - توضيح توزيع هذه المؤسسات تبعاً لموقعها برسم بياني :

٨ - كانت الارقام القياسية لنفقة المعيشة ولاسعار المجلة في احدى الدول في الفترة : ١٩٥٢ - ١٩٦١ كالاتي (على أساس ١٩٣٩) :

السنوات	الرقم القياسي لنفقة المعيشة	الرقم القياسي لأسعار المجلة
١٩٥٢	٣١٧	٣٧٢
١٩٥٣	٢٩٦	٣٥٥
١٩٥٤	٢٨٤	٣٤٥
١٩٥٥	٢٨٣	٣٥١
١٩٥٦	٢٩٠	٣٨٩
١٩٥٧	٣٠٢	٤٢٢
١٩٥٨	٣٠٢	٤١٧
١٩٥٩	٣٠٣	٤١٧
١٩٦٠	٣٠٤	٤١٨
١٩٦١	٣٠٦	٤٢٥

وضح التغير في هذه الارقام القياسية في رسم بياني واحد .

٩ - البيانات الآتية عن عدد السكان وعدد الاسرة في احدى الدول في الفترة بين ١٩٥٢ - ١٩٦٢ :

السنة	عدد السكان بالآلف	عدد الأسرة
١٩٥٢	٢١٤٧٣	٢٤٠٩٣
١٩٥٣	٢٢٠٠٣	٢٤٩٦٨
١٩٥٤	٢٢٥٥٧	٢٥٨٩٢
١٩٥٥	٢٣٠٦٣	٢٦٩٨٤
١٩٥٦	٢٣٦٤٣	٢٣٣٧٢
١٩٥٧	٢٤٢١٧	٣٠٤٠٧
١٩٥٨	٢٤٧٩١	٣٢١١٢
١٩٥٩	٢٥٣٦٥	٣٣١٥٣
١٩٦٠	٢٦٠٦٥	٣٧٠٩٤
١٩٦١	٢٦٢٧٤	٣٧٨٨٦
١٩٦٢	٢٧٠٦٤	٣٧٩٤٢

قارن التغير في عدد السكان وعدد الأسرة برسم بياني .

١٠ - البيانات الآتية تبين انتاج بعض المصنوعات في احدى الدول في عامي ١٩٥٢ ، ١٩٦٠ :

١٩٥٢	١٩٦٠
عزل القطن — ٥٦ الف طن	١٠٥ الف طن
منسوجات قطنية — ٢٢٠ مليون متر مربع	٤٨٢ مليون متر مربع
زيت بذرة القطن — ١٠٦ الف طن	١٠٤ الف طن
نبيذ — ١٦ الف طن	٣,٥ الف طن
سكر — ١٨٨ الف طن	٢٣٢ الف طن
اسمنت — ٩٤٧ الف طن	٢٠٤٧ الف طن

وضح التغير النسبي في الانتاج بين هذين العامين برسم بياني . احسب نسب

التغير على أساس عام ١٩٥٢ .

١١ - الآتي ارقام قياسية لأسعار المجلة في دولة ما في خلال عام ١٩٥٨ على أساس ١٩٤٠ .

الشهر : يناير - فبراير - مارس - أبريل - مايو - يونيو
الرقم القياسي : ١٠٩ ٩٨ ١٠٥ ٩٦ ٩٤ ٩٣

الشهر : يوليو - أغسطس - سبتمبر - أكتوبر - نوفمبر - ديسمبر
الرقم القياسي : ٩٣ ٩٧ ٩٦ ١٩٤ ١٠٣ ١٠٨

وضح التغير في هذه الأرقام .

١٢ - الآتي بيانات عن مجلة قروض حكومة دولة ما تبعاً لأنواعها المختلفة (الأرقام بملايين الليرات) .

السنة	قروض داخلية طويلة الأجل	قروض داخلية قصيرة الأجل	قروض أجنبية
١٩٤٥	٦١٠٠	٨٠٠	١٠٠٠
١٩٤٦	٦٢٠٠	٧٠٠	١٠٠٠
١٩٤٧	٦٣٠٠	٨٠٠	١٠٠٠
١٩٤٨	٦٤٠٠	٩٠٠	١٠٠٠
١٩٤٩	٦٦٠٠	١٥٠٠	١٢٠٠
١٩٥٠	٧٧٠٠	٢٨٠٠	١٣٠٠
١٩٥١	٩٨٠٠	٣٣٠٠	١٥٠٠
١٩٥٢	١١٦٠٠	٤١٠٠	١٦٠٠
١٩٥٣	١٣٥٠٠	٥٠٠٠	١٧٠٠

وضح التغير الذي حدث في تكوين قروض هذه الحكومة برسم بياني .

١٣ - في دراسة عن المساكن ، استعملت الاستمارة الآتية :

- ١ - رقم المسكن المسلسل _____
- ٢ - موقع المسكن المحافظة _____ القضاء _____
- ٣ - عدد الأسر التي تسكن في المسكن _____
- ٤ - نوع بناء السكن طين _____ طوب _____
اسمنت _____ خيام _____
- ٥ - عدد الغرف _____
- ٦ - هل السكن ملك _____ ايجار _____ هبة _____
- ٧ - مصدر المياه النقية مياه جارية _____ حنفية عامة _____
بئر خاض _____ بئر عام _____
- ٨ - اذا كان المسكن بالايجار كم الايجار السنوي _____
- ٩ - عدد الافراد الذين يسكنون المسكن _____ منهم
ذكور و _____ اناث _____
منهم _____ فوق سن ١٣
و _____ بين ٥ سنوات و ١٣ سنة
و _____ أقل من ٥ سنوات
- ١٠ - طريقة تصريف القاذورات نجاري عامة _____
بئر _____ طريقة أخرى _____
- ١١ - نوع سقف البناء اسمنت _____ خشب _____
مواد أخرى _____

- ١٢ - مساحة المسكن _____ متر مربع
- ١٣ - هل هناك حديقة للمسكن نعم _____ لا _____
- ١٤ - متى اسس المسكن عام _____

المطلوب

- ١ - تصميم الدليل لترميز البيانات السابقة
- ٢ - تصميم الكشف الذي يمكن أن يستخدم لتوقيع الدليل عليه
- ٣ - افترض مسكن له اجابات معينة على هذه الاسئلة ورمز هذه الاجابات في كشف الترميز الذي صممه .
- ١٤ - قامت دولة ما باجراء تعداد صناعي وكانت استمارة التعداد تحتوي على البيانات الآتية :

- | | |
|------------------------------|----------|
| ١ - اسم المؤسسة | |
| ٢ - موقع المؤسسة | القضاء |
| ٣ - نشاط المؤسسة | المحافظة |
| ٤ - الكيان القانوني للمؤسسة | |
| ٥ - عدد العمال الذكور | |
| ٦ - عدد العمال الاناث | |
| ٧ - مجموع عدد العمال | |
| ٨ - مجموع أجور الذكور في شهر | |
| ٩ - مجموع أجور الاناث في شهر | |
| ١٠ - المجموع العام للاجور | |
| ١١ - قيمة الانتاج | |

- ١٢ - قيمة المواد الخام الداخلة في الانتاج
 ١٣ - أنواع الآلات المستخدمة في الانتاج
 ١٤ - أنواع مولدات الطاقة المستخدمة
 ١٥ - قيمة الاستثمار الكلي في عام .

المطلوب

- ١ - اعداد الدليل الخاص بهذه البيانات
 ٢ - تصميم كشف توقع فيه أرقام الدليل ويجري التثقيب على أساسه .
 ١٥ - فيما يلي استمارة تشخيص خاصة بمرض البلهارسيا في أحد المستشفيات

اسم المريض رقم الادخال
 عنوان المريض تاريخ الادخال
 تاريخ الخروج
 مدة البقاء في المستشفى

الجنس ذكر _____ انثى _____
 العمر السنين _____ الأشهر _____
 المستوى الاقتصادي والاجتماعي
 عالي _____ وسط _____ فقير _____ لم يبين _____
 الديانة مسلم _____ مسيحي _____ ديانة اخرى _____ لم يبين _____
 الوزن عند الدخول _____ كيلو _____ جرام
 الوزن عند الخروج _____ كيلو _____ جرام
 الطول _____ سم

نوع مكان السكن _____ مدينة _____ قرية _____ بدو _____
 مدينة قروية _____ لم يبين _____
 المرض الأساسي أو الايذاء الذي ادى للدخال _____
 أحوال مرضية أخرى ساعدت (حادة أو مزمنة) _____

التشخيص

الاسنان	نعم	لا	لم يبين
كبد متمد	نعم	لا	لم يبين
بنكرياس متمد	نعم	لا	لم يبين
حمى	نعم	لا	لم يبين
اسهال حاد	نعم	لا	لم يبين
اسهال مفاود	نعم	لا	لم يبين
اسهال مع صديد	نعم	لا	لم يبين
اسهال مع دم	نعم	لا	لم يبين
اسهال مع نقط دهنية	نعم	لا	لم يبين
اسهال غير محدد	نعم	لا	لم يبين
جلد مقشر	نعم	لا	لم يبين

سبب الخروج من المستشفى :

العلاج انتهى _____ خروج بدون اذن _____ حول الى مستشفى _____
 آخر _____ الموت _____ سبب آخر _____
 لم يبين _____

الحالة عند الخروج .

تحسن _____ لم يتغير _____ أسوأ _____ وفاة _____
لم يبين _____

المطلوب :

- ١ - تصميم الدليل الحاض بترميز البيانات السابقة .
- ٢ - تصميم بطاقة التثقيب الخاصة بهذا الدليل .
- ٣ - تصميم كشف للترميز مع مثال .

١٦ - الأرقام الآتية تبين أطوال وأوزان مجموعة من الطلبة :

الطول مم	الوزن كيلو	الطول مم	الوزن كيلو	الطول مم	الوزن كيلو
١٧٠	٧٥	١٥٥	٦٠	١٥٤	٥٧
١٥٥	٦٠	١٤١	٥٢	١٦٨	٧٥
١٥٦	٦٣	١٤٧	٥٥	١٦٩	٨٢
١٤٣	٥٥	١٦٦	٧٠	١٥٥	٦٥
١٤٩	٥٩	١٦٩	٧٩	١٦٣	٦٣
١٦٨	٧٨	١٥٢	٥٤	١٥٤	٥٩
١٤٦	٦٢	١٧١	٨١	١٦٩	٧٦
١٥٣	٥٧	١٦٣	٥٩	١٧٣	٧٨
١٦٤	٦٢	١٧٣	٧٩	١٤٤	٥٥
١٦٩	٧٩	١٤٨	٥٧	١٦٥	٧٧
١٤٨	٥١	١٥٢	٥٥	١٦٩	٧٤
١٦٦	٦٦	١٧٤	٨٠	١٤٨	٥٧
١٥٤	٥٣	١٦٦	٧٥	١٦٣	٦٩
١٦٤	٦٩	١٥٣	٦٤	١٧١	٧٣
١٤٤	٥١	١٦٧	٧٤	١٦٣	٦٨
١٤٠	٥٩	١٦٨	٧٦	١٦٦	٧٥
١٤٢	٥٨	١٤٣	٥٠	١٥٧	٦٤
١٥٦	٦١	١٦١	٦٨	١٥٩	٦٦
١٦٣	٦٧	١٧٤	٧٨	١٥٦	٦٧
١٦٥	٦٩	١٧٣	٧٦	١٤٣	٥٩

٢ - المطلوب ترتيب هذه البيانات في جدول تكراري مزدوج .

١٧ - الآتي توزيع أوزان مجموعة من الطلبة :

الوزن بالكيلو	عدد الطلبة
٤٥	٢
٥٠	٣
٥٥	٣
٦٠	٦
٦٥	١٢
٧٠	٢٣
٧٥	٣١
٨٠	١٨
٨٥	١٨
٩٠	١٠
٩٥	٦

المطلوب :

- ١ - توضيح هذا التوزيع بهستو جرام .
- ٢ - توضيح هذا التوزيع بمضلع تكراري .
- ٣ - توضيح هذا التوزيع بمنحنى تكراري .
- ٤ - توضيح هذا التوزيع بمنحنى متجمع صاعد وهابط .
- ٥ - تحديد عدد الطلبة الذين يقل وزنهم عن ٨٦ كيلو .
- ٦ - تحديد الوزن الذي يقل عنه نصف أفراد المجموعة .
- ٧ - تحديد الوزن الذي يقل عنه ربع أفراد المجموعة .
- ٨ - تحديد عدد الطلبة الذين تتراوح أوزانهم بين ٦٢ و ٦٨ كيلو .
- ٩ - تحديد عدد الطلبة الذين تكون أوزانهم ٧٢ كيلو وأكثر .

١ الوسط الحسابي والهندسي والتوافقي والمتوال والوسط

الفصل الخامس

المتوسطات

بعد تلخيص البيانات في توزيع تكراري (بالنسبة للظواهر القيمية المتغيرة) ننتقل إلى التحليل الاحصائي لهذا التوزيع الذي يتكون أساساً من ثلاث أبحاث هي :

١ - البحث عن القيمة التي تتركز حولها القيم التي تتخذها الظاهرة في تغيرها من وحدة الى اخرى (المتوسط) .

٢ - البحث عن مدى تشتت القيم حول هذه القيمة المتوسطة (التشتت) .

٣ - البحث عن تماثل التوزيع حول القيمة المتوسطة (الالتواء) .

وفي هذا الفصل سنناقش مقاييس التركز التي جرى العرف الاحصائي على تسميتها بالمتوسطات وهي تشمل الوسط الحسابي والهندسي والتوافقي والمتوال والوسيط وغير ذلك من المتوسطات التي لن نتعرض لها في هذا الكتاب .

والفكرة الاساسية التي يقوم عليها موضوع المتوسطات هي أن تمثيل التوزيع التكراري بقيمة واحدة يبرره ميل المجموعات الكبيرة من الوحدات نحو أن تتركز قيمها حول قيمة معينة تنحرف عنها القيم الاخرى بشيء من

$$س١ + س٢ + س٣ + س٤ + س٥ + س٦ + س٧ + س٨ + س٩ + س١٠$$

$$س١ + س٢ + س٣ + س٤ + س٥ + س٦ + س٧ + س٨ + س٩ + س١٠$$

الانتظام ، هذه القيمة هي التي نسميها بالمتوسط وان كانت تتخذ أسماء مختلفة تبعاً للأساس الذي يبنى عليه قياسها . وبذلك يكون أي متوسط هو القيمة التي نعتبرها ممثلة لمجموعة القيم التي حسب لها هذا المتوسط . وإذا كانت الأنواع المختلفة من المتوسطات تعتبر ممثلة للقيم التي حسبت لها مع اختلاف قيمة كل نوع منها، فليس في ذلك أي تناقض حيث أن هناك أسباب معينة خاصة بكل متوسط والتي على أساسها يعتبر ممثلاً للقيم التي حسب لها وسوف ندرس هذه الأسباب عندما نأتي إلى مناقشة كل متوسط على حدة .

الوسط الحسابي :

نفترض ان لدينا ظاهرة معينة س تتغير قيمتها من وحدة الى اخرى (مثلا أجور مجموعة من العمال) وبذلك يصبح لدينا القيم س١ ، س٢ ، س٣ ، س٤ ، سن ، فإن كان عدد هذه القيم (عدد الوحدات موضوع الدرس) ن يكون الوسط الحسابي لهذه القيم :

$$\bar{س} = \frac{س١ + س٢ + س٣ + س٤ + + سن}{ن}$$

اي ان س- = $\frac{\text{مجموع القيم}}{ن}$ (الرمز مح يدل على مجموع القيم)

وبذلك يكون مح س = ن س-
 عدد القيم × المتوسط = مجموع

ما هي خائص هذه القيمة س- حق نعتبرها ممثلة للقيم التي حسبت لها ؟ ان عملية حساب هذه القيمة تجعلها توازن بين القيم الصغيرة والقيم الكبيرة من حيث انحرافها عنها ؛ بمعنى ان مجموع الانحرافات القيم الكبيرة عن س- = مجموع الانحرافات القيم الصغيرة عن س- ، أي ان المجموع الجبري لانحرافات جميع القيم عن الوسط الحسابي = صفر ولايثبات ذلك نحسب الانحرافات القيم عن س- ونجمعها كالآتي :

① مجموع الانحرافات للقيم الكبيرة عن المتوسط
 ٢٢٨ انحرافات القيم الصغيرة عن المتوسط
 المجموع الجبري لانحرافات جميع القيم عن الوسط
 صفر

المتوسط الحسابي و الوسط الفرضي + $\frac{\sum x}{n}$ مجموع الانحرافات عن الوسط الفرضي.

وبالقسمة على n يكون $-س - و =$

فاذا رمزنا لمجموع الانحرافات عن الوسط الفرضي بالرمز مع ح

$$\frac{\sum x}{n} + و = -س$$

المعكوزة و $\frac{\sum x}{n} + و = -س$

الوسط الحسابي المرجح :

نلاحظ انه في المثال السابق كانت القيم متعادلة في الاهمية ، لكن نفترض أن القيمة س، تكررت ك، مرة والقيمة س، تكررت ك، مرة والقيمة س، تكررت ك، مرة وهكذا ، وان ك، و ك، و ك، غير متساوية هل يحسب الوسط الحسابي بنفس الطريقة السابقة . لا شك ان استعمال نفس الطريقة السابقة يجعلنا نحصل على قيمة مضللة بعض الشيء اذ انه يجب أن يؤخذ عدد مرات تكرار كل قيمة في حسابنا حتى نحصل على وسط حسابي غير مضلل . ان عدد مرات تكرار القيم يدل على اهميتها ولذلك نسميها بالاوزان والوسط الحسابي لهذا النوع من القيم نسميه الوسط الحسابي المرجح حيث اننا نرجح كل قيمة بوزنها الذي يدل على اهميتها ويكون ذلك كالآتي :

$$\frac{س_1 \times ك_1 + س_2 \times ك_2 + س_3 \times ك_3 + س_4 \times ك_4}{ك_1 + ك_2 + ك_3 + ك_4} = -س$$

والوسط المحسوب بهذه الطريقة يتوفر فيه كذلك الخاصيتان التي سبق الإشارة اليها ، حيث اننا في الواقع لم نغير شيئاً في طريقة الحساب اذ بدل أن نكرر س، جمعاً ك، مرة ضربناها في ك، وكذلك بدل أن نكرر س،

جما ك_٢ مرة ضربناها في ك_٢ ، أما المقام ك_١ + ك_٢ + ك_٣ + ك_٤ فهذا
ايضاً يدل على العدد الكلي للقيم الموجودة في البسط .

والآن نحسب الانحرافات القيم عن وسطها الحسابي ونجمعها كالآتي :

$$(١٠ - ١٠) + (١٠ - ١٠) + (١٠ - ١٠) + (١٠ - ١٠) + (١٠ - ١٠) + (١٠ - ١٠) + (١٠ - ١٠) + (١٠ - ١٠) + (١٠ - ١٠) + (١٠ - ١٠)$$

$$= ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠$$

$$= ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠$$

$$= ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠ - ١٠$$

حيث ان س_١ ك_١ = مجموع القيمة الاولى وس_٢ ك_٢ = مجموع القيمة
الثانية وس_٣ ك_٣ = مجموع القيمة الثالثة ، وبذلك يكون مجموعها جميعاً
مساوياً للمجموع الكلي للقيم الظاهرة س .

مثال ١ :

نفترض ان لدينا خمس عمال أجورهم كالآتي ٨ ، ٣ ، ٥ ، ٩ ، ١٠

$$\bar{x} = \frac{٣٥}{٥} = ٧$$

نحسب الآن الانحرافات القيم عن ٧ ونجمعها :

$$١ + ٤ - ٤ + ٢ + ٣ = ٦$$

نربع هذه الانحرافات نحصل على ١ + ١٦ + ٤ + ٤ + ٩ = ٣٤

إذا أخذنا قيمة أخرى فرضية ٦ ونحسب الانحرافات عنها

$$٢ + ٣ - ١ + ٣ + ٤ = ١١$$

نربع هذه الانحرافات نحصل على ٤ + ٩ + ١ + ٩ + ١٦ = ٣٩

وهو أكبر من مجموع مربع الانحرافات عن الوسط الحسابي .
 يمكننا أن نحصل على الوسط الحسابي عن طريق هذه القيمة الفرضية ٦
 كالآتي : -

$$س = -و + \frac{م ح}{ن}$$

الوسط الحسابي ، الوسط الفرضي

$$٧ = \frac{٥}{٥} + ٦ =$$

مثال ٢ :

نفترض أن ١٠٠٠ عامل يحصل كل منهم على أجر ٥ ليرات .
 وان ١٠ عمال يحصل كل منهم على أجر ١٠٦ ليرات .

إذا حسبنا المتوسط كالاتي $\frac{٥ + ١٠٦}{٢} = ٥٥,٥$ نلاحظ انه قيمة مضللة

حيث أن معظم العمال يتقاضون في الواقع أجراً لا يزيد عن ٥ ليرات .

ولكن بحساب الوسط المرجح

$$٦ = \frac{١٠٦٠}{١٠١٠} = \frac{١٠٦ \times ١٠ + ٥ \times ١٠٠٠}{١٠١٠}$$

وهو أجر قريب جداً من أجر غالبية العمال :

نحسب انحرافات القيم من المتوسط المرجح ٦ ونجمع هذه الانحرافات :

$$(-) ١٠٠٠ + ١٠ (١٠٠) = صفر .$$

إذا أخذنا قيمة قيمة أخرى فرضية ٩، ونحسب الانحراف عنها :

$$\begin{array}{r} ١٠٠٠ \times ١ - \\ ١٠ \times ١٠٠ \end{array}$$

الوسط الحسابي من - ٩ = ٣٠٣٠ - ١٠١٠

$$3030 - = 10(97) + 1000 \times (4-)$$

الوسط الحسابي من - ٩ = و + $\frac{\text{مجموع}}{ن}$

$$\frac{3030-}{1010} + 9 =$$

$$6 = 3 - 9 =$$

الوسط الحسابي لتوزيع تكراري :

لحساب الوسط الحسابي لتوزيع تكراري يمكن أن تتبع الخطوات الآتية :

١ - نحسب مركز كل فئة حيث أن القيمة المتوسطة للفئة .

٢ - نضرب مركز كل فئة في تكرارها ونجمع النتائج .

٣ - نستخدم القانون من - $\frac{\text{مجموع ك}}{\text{مجموع}}$ (نلاحظ ان هذه الطريقة لا

لا تختلف عن طريقة الوسط الحسابي المرجح التي سبق الاشارة اليها وليس هناك من فرق الا اننا عبرنا عن القيم بالرمز م بدلا من الرمز س) .

مثال ٣ : (توزيع تكراري غير متصل)

الفئات	التكرارات (ك)	المراكز (م)	المراكز التكرارات (م ك)
٦٢-٦٠	٥	٦١ ✓	٣٠٥
٦٥-٦٣	١٨	٦٤ ✓	١١٥٢
٦٨-٦٦	٤٢	٦٧ ✓	٢٨١٤
٧١-٦٩	٢٧	٧٠ ✓	١٧٨٠
٧٤-٧٢	٨	٧٣ ✓	٥٨٤
	١٠٠		٦٧٤٥

٢٣٣

$$\frac{233}{100} + 70$$

١- نكتب كل فئة
٢- نضرب مركز كل فئة في تكرارها ونجمع النتائج
٣- نستخدم القانون من - $\frac{\text{مجموع ك}}{\text{مجموع}}$ (نلاحظ ان هذه الطريقة لا تختلف عن طريقة الوسط الحسابي المرجح التي سبق الاشارة اليها وليس هناك من فرق الا اننا عبرنا عن القيم بالرمز م بدلا من الرمز س)

$$\text{س-} = \frac{6745}{100} = 67.45$$

(نلاحظ اننا حسبنا مركز الفئة الأولى ٦١ وأضفنا باستمرار ٣ الذي هو مدى الفئة كي نحصل على مراكز الفئات التالية) .

مثال ٤ : (توزيع تكراري متصل)

ف	ك	م	م × ك
٦٠ -	٥	٦١.٥	٣٠٧.٥
٦٣ -	١٨	٦٤.٥	١٩٦١
٦٦ -	٤٢	٦٧.٥	٢٨٣٥
٦٩ -	٢٧	٧٠.٥	١٩٠٣.٥
٧٢ -	٨	٧٣.٥	٥٨٨
	<u>١٠٠</u>		<u>٦٧٩٥</u>

$$\text{س-} = \frac{6795}{100} = 67.95$$

مثال ٥ : (توزيع تكراري غير منتظم)

ف	ك	م	م × ك
٥ -	٥	٧.٥	٣٧.٥
١٠ -	١٨	١٥	٢٧٠
٢٠ -	٤٢	٢٥	١٠٥٠
٣٠ -	٢٧	٤٠	١٠٨٠
٥٠ -	٨	٦٠	٤٨٠
	<u>١٠٠</u>		<u>٢٩١٧.٥</u>

$$س- = \frac{٢٩١٧,٥}{١٠٠} = ٢٩,١٧٥$$

ان استخدام الطريقة السابقة س- = $\frac{م ك}{ح ك}$ يستدعي عمليات

حسابية طويلة يضيع فيها الوقت ولذلك يكون من الأفضل أن نلجأ إلى الطريقة الثانية التي يحسب فيها الوسط بناء على قيمة نفترضها وبذلك نستعمل القانون :

$$س- = و + \frac{م ح ك}{ح ك}$$

ونلاحظ ان هذا القانون لا يختلف عن القانون الذي سبق الاشارة اليه فالقيمة م ح ك تدل على مجموع الانحرافات عن الوسط الفرضي وادخلنا ك في الحساب لأن الانحرافات القيم عن الوسط الفرضي تتكرر تبعاً لتكرار القيمة نفسها . كذلك يكون م ك هو نفسه عبارة عن عدد القيم .

مثال ٦ :

ف	ك	م	ح	ح ك
٦٢ - ٦٠	٥	٦١	- ٦	- ٣٠
٦٥ - ٦٣	١٨	٦٤	- ٣	- ٥٤
٦٨ - ٦٦	٤٢	٦٧	صفر	
٧١ - ٦٩	٢٧	٧٠	+ ٣	+ ٨١
٧٤ - ٧٢	٨	٧٣	+ ٦	+ ٤٨
	<u>١٠٠</u>			<u>٤٥</u>

٢٣٥

$$\frac{٤٥}{١٠٠} + ٦٧ = س-$$

الوسط الفرضي + الوسط الفعلي = مجموع ل
 طول ل

$$\therefore -س = ٦٧ + \frac{٤٥}{١٠٠} \times ٦٧,٤٥$$

ونلاحظ ان الانحرافات ٦ - ، ٣ + ، ٣ + ، ٦ + تختصر جميعاً على طول الفئة ٣ ولذلك تسهلاً للعمل تختصر هذه الانحرافات وفي هذه الحالة نستخدم القانون :

$$س = -و + \frac{\text{مجموع ك}}{\text{مجموع ح}} \times \text{طول الفئة} + ٦٧ \times \frac{٤٥}{١٠٠}$$

نلاحظ اننا ضرب مجموع ك \times طول الفئة حيث اننا عند ايجاد مجموع الانحرافات عن الوسط الفرضي اختصرناها جميعاً على طول الفئة ولذلك نعود فنضربها في طول الفئة حتى نرجعها الى اصلها ، أما إذا كانت الفئات غير منتظمة نحسب الانحرافات الواقعية عن الوسط الفرضي ولا يكون هناك داعي للضرب في طول الفئة عند تطبيق القانون .

مثال ٧ :
 توزيع سنو
 الانحرافات
 لواقعية
 الفرضي
 مثال
 طول الفئة عند تطبيق
 القانون

ف	ك	م	ح	ح ك
٦٢ - ٦٠	٥	٦١	٢ -	١٠ -
٦٥ - ٦٣	١٨	٦٤	١ -	١٨ -
٦٨ - ٦٦	٤٢	٦٧	صفر	صفر
٧١ - ٦٩	٢٧	٧٠	١ +	٢٧ +
٧٤ - ٧٢	٨	٧٣	٢ +	١٨ +
	١٠٠			١٥

$$\therefore \text{س} = 67 + 3 \times \frac{10}{100} = 67.3$$

الوسيط :

طريقة الوسط يجب ترتيب القيم ترتيباً تصاعدياً
عندها إذا حسبناه عن الوسط تكون القيمة التي تقع في الوسط أي التي يقل
الوسيط لأي مجموعة من القيم هو القيمة التي تقع في الوسط أي التي يقل
عنها عدد من القيم يساوي عدد القيم التي تزيد عنها ، وهذه هي الخاصية التي
على أساسها يكون الوسيط هو القيمة الممثلة لمجموعة القيم التي حسب لها .
فالوسيط هو القيمة التي توازن بين القيم الصغيرة والقيم الكبيرة من حيث ① بعد ترتيب
عددها .

ولحساب الوسيط يجب ترتيب القيم ترتيباً تصاعدياً حتى إذا حسبناه في القيمة المطلوبة
المنتصف تكون القيم السابقة له أقل منه والقيم اللاحقة له أكبر منه . بعد الترتيب
ترتيب القيم نجد ترتيب الوسط ثم نبحث عن القيمة المقابلة لهذا الترتيب ، وإذا
كان ترتيب الوسيط يقع بين قيمتين نجد الوسط الحسابي لها .

مثال ٨ :

نفترض ان لدينا القيم ٧٨ ، ٧٢ ، ٦٨ ، ٨٧ ، ٨٤ ، ٩١ ، ونريد أن
نحسب لها الوسيط . نرتب القيم ترتيباً تصاعدياً كالآتي :

$$91, 87, 84, 78, 72, 68$$

$$\text{نحسب بعد ذلك ترتيب الوسيط وهو يساوي} = \frac{\text{عدد القيم} + 1}{2}$$

$$= \frac{1 + 6}{2} = 3.5 \text{ أي ان الوسيط هو القيمة الثالثة والنصف وهذا الترتيب يقع}$$

$$\text{بين } 78 \text{ ، } 84 .$$

$$\frac{78 + 84}{2} = 81$$

$$\checkmark \quad 81 = \frac{162}{2} = \frac{81 + 78}{2} = \text{الوسيط} \therefore$$

وعندما تكون القيم مبنوبة في توزيع تكراري يجب اولا وضع التوزيع في شكل متجمع صاعداً أو هابطاً (يكون ذلك ترتيباً للقيم) ثم نحدد ترتيب الوسيط وهو يساوي مجموع التكرارات مقسوماً على ٢ ثم نحاول ان نجد القيمة المقابلة لهذا الترتيب ولتوضيح ذلك نفترض المثال الآتي :

مثال ٩ :

التكرار المتجمع الصاعد	الحدود العليا للفئات	ك	ف
٣	أقل من ١٢٧	٣	١١٨ -
٨	١٣٦ »	٥	١٢٧ -
١٧	١٤٥ »	٩	١٣٦ -
٢٩	١٥٤ »	١٢	١٤٥ -
٣٤	١٦٣ »	٥	١٥٤ -
٣٨	١٧٢ »	٤	١٦٣ -
٤٠	١٨١ »	٢	١٨١ - ١٧٢
		٤٠	

بين التكرار الصاعد ١٧ والتكرار
بط وهي للقيمة المقابلة لهذا الترتيب تقع
تغير في التوزيع هو تغير منتظم يكون
مستقيماً ويكون ميله بين هاتين القيمتين

$$\frac{15 - 29}{15 - 145}$$

$$\frac{15 - 29}{15 - 145}$$



وحيث انه خط مستقيم يكون هذا الميل مساويا للميل عند $\frac{15 - 29}{145 - 15}$ النقطة التي تمثل ترتيب الوسيط ٢٠

$$\frac{15 - 29}{145 - 15}$$

اي : $\frac{15 - 20}{145 - س}$ (س تدل على قيمة الوسيط)

$$\frac{15 - 29}{145 - س} = \frac{12 - 27}{145 - س} \therefore \frac{3}{145 - س} = \frac{12}{9} \therefore 27 = 1740 = س \cdot 12 \therefore س = 12$$

$$1767 = س \cdot 12$$

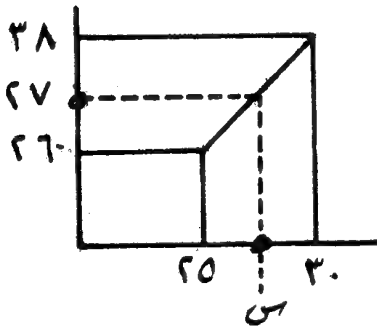
$$س (الوسيط) = 147,25$$

$$\frac{12}{9} = \frac{3}{145 - س}$$

مثال ١٠ :

ف	ك	الحدود العليا للفئات	التكرار المتجمع الصاعد	س - ١٧
١٠ -	٣	أقل من ١٥	٣	س - ١٤٥
١٥ -	٧	٢٠ د	١٠	
٢٠ -	١٦	٢٥ د	٢٦	
٢٥ -	١٢	٣٠ د	٣٨	
٣٠ -	٩	٣٥ د	٤٧	
٣٥ -	٥	٤٠ د	٥٢	
٤٠ - ٤٥	٢	٤٥ د	٥٤	
	٥٤			

ترتيب الوسيط = $\frac{54}{2} = 27$ ويقع بين التكرار المتجمع الصاعد ٢٦ والتكرار الصاعد ٣٨ ، وتكون قيمة الوسيط هي القيمة المقابلة لهذا الترتيب وهي تقع بين ٢٥ ، ٣٠ .



$$\frac{26 - 27}{25 - س} = \frac{26 - 38}{25 - 30} \therefore$$

$$\frac{1}{25 - س} = \frac{12}{5} \therefore$$

$$5 = 300 - س \quad \therefore$$

$$305 = س \quad \therefore$$

$$\therefore س (الوسيط) = \frac{305}{12} = 25 ر 41$$

ونلاحظ ان قيمة الوسيط تقع في الفئة ٢٥ - ٣٠ حيث ان عدد للتكرارات السابقة لهذه الفئة $26 = 16 + 7 + 3$ أما ترتيب الوسيط فهو ٢٧ وبذلك تكون قيمته واقعة في هذه الفئة وحتى تحصل على هذه القيمة نحسب ما يقابل الزيادة في ترتيب الوسيط عن مجموع التكرارات السابقة (٢٦ - ٢٧) من طول الفئة ٥ .

$$\text{وذلك يساوي } 0.42 = 5 \times \frac{1}{12} \therefore \text{الوسيط} = 25 +$$

$$0.42 = 25.42$$

الرابع الأول هو القيمة التي يسبقها ربع عدد القيم ويزيد عنها
 شئ أو أربع القيم ، الرابع الثالث فهو القيمة التي يسبقها

ويمكن استخدام الطرق السابقة في إيجاد جميع القيم الموضعية الأخرى مثل
 مثل الربع الأدنى أو الربع الأول والربع الأعلى أو الثالث والعشر الأول عدد القيم
 والعشر الثاني وهكذا . والربع الأول هو القيمة التي يسبقها ربع عدد القيم
 ويزيد عنها ثلاث أرباع عدد القيم . أما الربع الثالث فهو القيمة التي يسبقها
 ثلاث أرباع عدد القيم ويزيد عنها ربع عدد القيم . والعشر الأول هو القيمة
 التي يقل عنها عشر عدد القيم ويزيد عنها تسع أعشار عدد القيم .

مثال ١١ :

العشر الأول هو القيمة
 التي يقل عنها عشر
 عنها تسع أعشار
 القيم

ف	ك	الحدود العليا للفئات	التكرار المتجمع الصاعد
٣٠ -	١	أقل من ٤٠	١
٤٠ -	٣	أقل من ٥٠	٤
٥٠ -	١١	أقل من ٦٠	١٥
٦٠ -	٢١	أقل من ٧٠	٣٦
٧٠ -	٤٣	أقل من ٨٠	٧٩
٨٠ -	٣٢	أقل من ٩٠	١١١
٩٠ - ١٠٠	٩	أقل من ١٠٠	١٢٠

ترتيب الوسيط = $\frac{120}{2} = 60$ واقع بين ٣٦ ، ٧٩

$$\frac{36 - 60}{70 - 60} = \frac{36 - 69}{70 - 80} \therefore$$

$$\frac{24}{70 - 60} = \frac{43}{10} \therefore$$

$$240 = 1010 - \text{س } 43 \therefore$$

$$3250 = \text{س } 43 \therefore$$

$$75,5 = \text{س } \therefore$$

$$\text{ترتيب الربع الاعلى} = \frac{120}{4} = 30 \text{ واقع بين } 15, 36$$

$$\frac{15 - 30}{60 - \text{س}} = \frac{15 - 36}{60 - 70} \therefore$$

$$\frac{15}{60 - \text{س}} = \frac{21}{10} \therefore$$

$$140 = 1260 - \text{س } 21 \therefore$$

$$1410 = \text{س } 21 \therefore$$

$$671 = \text{س } \therefore$$

$$\text{ترتيب الربع الاعلى} : \frac{120}{4} \times 3 = 90 \text{ واقع بين } 79, 111$$

$$\frac{79 - 90}{80 - \text{س}} = \frac{79 - 111}{80 - 90} \therefore$$

$$110 = 2560 - \text{س } 32 \therefore \frac{11}{80 - \text{س}} = \frac{23}{10} \therefore$$

$$2670 = \text{س } 32 \therefore$$

$$834 = \text{س}$$

ويمكن تحديد قيمة الوسيط وغيره من القيم الموضعية بواسطة الرسم البياني ، وذلك برسم المنحنى المتجمع الصاعد أو الهابط من واقع التوزيع ، ثم نحدد موقع ترتيب الوسيط أو أي قيمة موضعية أخرى على المحور الرأسي ، ثم نصل النقطة الدالة على هذا الموقع مع المنحنى المتجمع ومن نقطة التقاطع نسقط عمود على المحور الأفقي يقابله في نقطة تكون القيمة عندها هي الوسيط أو القيمة الموضعية التي نبحث عنها . وإذا رسمنا المنحنى الصاعد والهابط في رسم واحد فإنها يتقاطعان في نقطة تحدد ترتيب الوسيط ومنها ننزل عمود على المحور الأفقي يقابله في قيمة تكون هي الوسيط .

المسئله :

المسئله لمجموعة من القيم هو القيمة الأكثر انتشاراً أو الأكثر تكراراً / بين القيم ، وبمعنى آخر هي القيمة الشائعة ، وهذا هو الأساس الذي بناء عليه يعتبر المسئله وسطاً ممثلاً للقيم التي حسب لأجلها . إلا أن هذه القيمة قد لا توجد وحق عند وجودها قد لا تكون قيمة ذات دلالة فعلية .

مسئله ١٢ :

للمجموعة القيم $١٨ ، ١٢ ، ١١ ، ١٠ ، ٩ ، ٩ ، ٧ ، ٥ ، ٢ ، ٢$ المسئله هو ٩ .

للمجموعة القيم : $١٦ ، ١٥ ، ١٢ ، ١٠ ، ٩ ، ٨ ، ٥ ، ٣$ لا يوجد مسئله .

للمجموعة القيم : $٩ ، ٧ ، ٧ ، ٧ ، ٥ ، ٥ ، ٤ ، ٤ ، ٤ ، ٣ ، ٢$ هناك مسئله ٧ .

وبالنسبة للتوزيع التكراري لا يكون من السهل تحديد قيمة المنوال . وسنناقش هنا طريقتين لإيجاد المنوال ولكن يجب أن نتذكر دائماً أن كلتا الطريقتين تقريبيتين . والخطوة الأولى في تحديد المنوال هي تحديد الفئة التي يكون تكرارها أكبر تكرار في التوزيع الموجود لدينا ، وواضح إذا غيرنا هذا التوزيع باستخدام فئات ذات مدى مختلف فسوف نحصل على فئة منوالية أخرى أي فئة يكون تكرارها هو الأكبر وبذلك تختلف قيمة المنوال بعض الشيء تبعاً لاختلاف مدى الفئة المستخدم في التوزيع ، ويعني ذلك أن قيمة المنوال ليست قيمة مستقرة وذلك هو السبب في عدم دقة قيمة المنوال كما سبق أن ذكرنا . وتنشأ هذه الصعوبة غالباً (صعوبة تحديد قيمة المنوال) بسبب تبويب البيانات في فئات واسعة المدى ، الأمر الذي يضيع معالم التوزيع ، أما إذا بوبنا البيانات في فئات ضيقة المدى فإن قيمة المنوال تستقر إلى حد ما عند رقم معين .

١١ الطريقة

لإيجاد المنوال

طريقة الحصول على قيمة دقيقة نوعاً ما للمنوال الذي يكون هو القيمة المقابلة لقيمة المنحنى ، حيث أن هذه القيمة سوف تمثل أكبر تكرار ، وذلك لأن تمهيد المنحنى يكون بمثابة وضع التوزيع في فئات ضيقة المدى . إلا أن هذه الطريقة بها ذات عيب يرجع إلى صعوبة تمهيد المنحنى التكراري تمهيداً دقيقاً خاصة وأن ذلك يتوقف إلى حد كبير على خبرة الباحث بهذا الموضوع .

الاستمرار في التكرار

والطريقة الحسابية الأولى لإيجاد المنوال تسمى طريقة الرافعة وتنحصر في تحديد الفئة المنوالية للتوزيع وهي الفئة ذات التكرار الأكبر ، كما قدمنا ، فيكون المنوال هو الحد الأدنى لهذه الفئة مضافاً إليه جزء من طول الفئة المنوالية يتناسب مع التكرارات السابقة واللاحقة لهذه الفئة والتي تعتبر كقوى مؤثرة على موضع المنوال فيها .

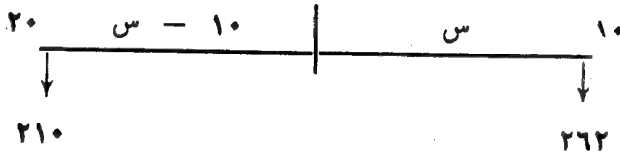
مثال ١٣ :

ك	ف
٢٦٢	صفر -
٣٣٣	١٠ -
٢١٠	٢٠ -
١٤٢	٤٠ - ٣٠

$$\frac{٢١٠}{٢٦٢ + ٢١٠} \times (١٠ - ٢٠) + ١٠ = \text{المنوال}$$

$$١٤٤٤٩ = ٤٤٤٩ + ١٠ =$$

ولتوضيح الاجابة نفترض ان الفئة المنوالية هي كرافعة يؤثر فيها القوتان ٢٦٢ ، ٢١٠ ويكون توازنها كالاتي :



$$\therefore ٢٦٢ \text{ س} = ٢١٠ (١٠ - \text{س})$$

$$\therefore ٢٦٢ \text{ س} = ٢١٠٠ - ٢١٠ \text{ س}$$

$$\therefore ٤٧٢ \text{ س} = ٢١٠٠$$

$$\text{س} = \frac{٢١٠٠}{٤٧٢} = ٤٤٤٩$$

$$\therefore \text{المنوال} = ١٠ + ٤٤٤٩ = ١٤٤٤٩$$

ويمكن وضع هذه الطريقة في قانون كالاتي :

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{التكرار اللاحق} - \text{التكرار السابق}}{\text{طول الفئة}}$$

والطريقة الثانية لحساب المنوال تسمى طريقة الفروق أو طريقة بيرسون تبعاً لاسم صاحبها وتقوم على أساس تحديد الفئة المنوالية كالطريقة السابقة تماماً ولكنها تختلف عنها في ان الذي يؤثر على موضع المنوال في هذه الفئة ليس هو التكرار السابق والتكرار اللاحق وإنما الفرق بين تكرار الفئة المنوالية وتكرار الفئتين السابفة واللاحقة لها . ويكون تحديد موضع المنوال في الفئة المنوالية على أساس تقسيمها بنسبة الفرق السابق الى الفرق اللاحق .

مثال ١٤ :

الفئة المنوالية	ك	ف
	٢٦٢	صفر ✓
	٣٣٣	١٠ ✓
	٢١٠	٢٠ ✓
	١٤٢	٤٠ - ٣٠ ✓

$$\text{والمنوال بهذه الطريقة} = ١٠ + \frac{٢٦٢ - ٣٣٣}{(٢١٠ - ٣٣٣) + (٢٦٢ - ٣٣٣)} \times ١٠$$

$$١٠ \times \frac{٧١}{١٢٣ + ٧١} + ١٠ =$$

$$١٣٦٦ = ٣٦٦ + ١٠ =$$

ولتوضيح الاجابة نقسم الفئة المنوالية تقسيماً جبرياً بنسبة (س) : ١٠ - س
وحسابياً بنسبة الفرق السابق ٧١ الى الفرق اللاحق ١٢٣ كالاتي :

$$\frac{س - ١٠}{١٢٣} = \frac{س}{٧١}$$

الفئة المنوالية = $\frac{س}{٧١}$

$$\frac{٧١}{١٢٣} = \frac{س}{س - ١٠} \therefore$$

الفرق السابق = $\frac{س}{س - ١٠}$

$$\therefore ١٢٣ س = ٧١ س - ٧١٠ \therefore$$

$$\therefore ١٩٤ س = ٨١٠$$

$$\therefore س = ٣٦٦$$

$$\therefore \text{المنوال} = ١٠ + ٣٦٦ = ٣٧٦$$

ويمكن وضع هذه الطريقة في صيغة قانون كالاتي :

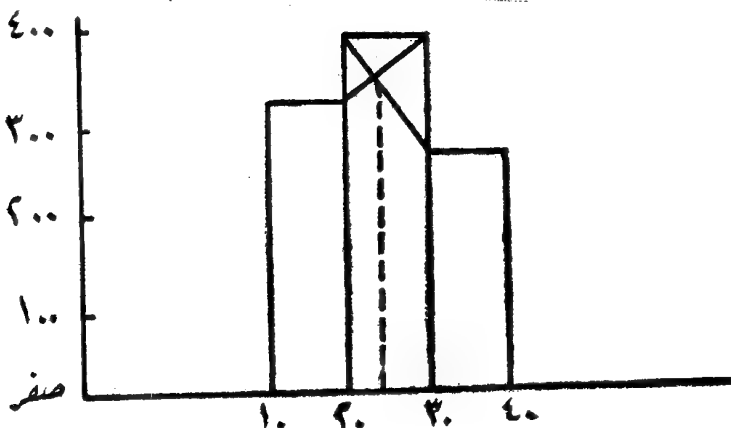
المنوال = الحد الأدنى للفئة المنوالية +

$$\times \text{طول الفئة} \frac{\text{الفرق السابق}}{\text{الفرق السابق} + \text{الفرق اللاحق}}$$

ونلاحظ ان المنوال محسوباً بالطريقة الاولى يختلف عن قيمته محسوباً بالطريقة الثانية ، وليس هذا من المستغرب حيث ان كلا من الطريقتين تقريبية كما سبق ان ذكرنا وخاصة اذا تذكرنا ان عدم الدقة يزداد كلما زاد مدى الفئة ، وفي هذا التوزيع مدى الفئة واسع نوعاً ما .

ويمكن التعبير عن هذه الطريقة بيانياً بالشكل الآتي وهو يوضح يحلله كيف تؤثر الفروق بين التكرارات السابقة وتكرار الفئة المنوالية على موضع المنوال في هذه الفئة ويمكن في الواقع استخدام هذه الطريقة البينائية والاستعاضة بها عن طريقة الفروق الحسابية .

هنا في ١٨/٤



ونلاحظ ان المنوال يقرب من الحد الادنى للفئة المنوالية كلما كان الفرق السابق أصغر من الفرق اللاحق ، والعكس يقرب المنوال من الحد الاعلى للفئة المنوالية كلما كان الفرق اللاحق أصغر من الفرق السابق ، فاذا تساوى الفرقان يكون المنوال هو مركز الفئة المنوالية ، وفي هذه الحالة تكون قيمة المنوال محسوبة بهذه الطريقة مساوية تماماً لقيمه محسوبة بطريقة الرافعة حيث يكون الفرقان متساويان عندما يكون التكرار السابق واللاحق متساويان وبذلك تتوازن الرافعة في منتصفها .

وفي بعض الحالات يكون تحديد المنوال أمراً معقداً لسبب وجود أكثر من قيمة واحدة تتركز عندها القيم . ومحدث ذلك إما لأن الفئات صغيرة المدى جداً بالنسبة الى عدد الوحدات التي تقوم بدراستها وبذلك يكون توزيعها في هذه الفئات لا يساعد على إظهار الفئة التي تتركز فيها القيم . وفي هذه الحالة يمكن التخلص من تعدد قيم المنحنى بتوسيع مدى الفئة شيئاً فشيئاً حتى نحصل على فئة منوالية واحدة ، واذا بقي التوزيع متعدد القيم بالرغم من ذلك فان هذا يكون دليلاً قاطعاً كما سبق ان ذكرنا على عدم تجانس البيانات التي نقوم بدراستها وفي هذه الحالة لا يكون للمنوال أي أهمية تذكر .

الوسط الهندسي :

الوسط الهندسي لأي مجموعة من القيم عددها n هو الجذر النوني لحاصل ضرب هذه القيم .

مثال ١٥ :

$$\frac{12, 10, 7, 6, 6, 5, 2}{12 \times 10 \times 7 \times 6 \times 6 \times 5 \times 2} \sqrt[7]$$

∴ لو الوسط الهندسي = $\frac{1}{7} (\text{لو } 2 + \text{لو } 5 + \text{لو } 6 + \text{لو } 6 + \text{لو } 7 + \text{لو } 10 + \text{لو } 12)$
(١٠ + ١٢)

$$0.8081 = 5667 \times \frac{1}{7} =$$

∴ الوسط الهندسي = ٦٤٣ الى أقرب رقمين عشريين .
واذا كانت القيم غير متساوية في الأهمية أي ان كلا منها له وزن معين يدل عليه عدد مرات تكرارها يكون الوسط الهندسي هو جذر مجموع الاوزان لحاصل ضرب القيم مرفوعة الى أوزانها .

مثال ١٦ :

القيم ٢ ، ٥ ، ٢٦ ، ٧
الاوزان ٢٥ ، ٣٨ ، ١٩ ، ١٨

$$\frac{25 \times 38 \times 19 \times 18}{25 \times 38 \times 19 \times 18} \sqrt[100]$$

$$\therefore \text{لو الوسط الهندسي} = \frac{1}{100} (25 \text{ لو } 2 + 28 \text{ لو } 5 + 19 \text{ لو } 6 + 28 \text{ لو } 7)$$

$$= \frac{1}{100} (25 \times 0.301 + 38 \times 6.99 + 19 \times 7.782 + 18 \times 0.8427)$$

$$= 63.0846 \times \frac{1}{100}$$

$$= 63.0846$$

$$\therefore \text{الوسط الهندسي} = 63.0846$$

واذا كانت القيم مبوبة في توزيع تكراري تتبع نفس الخطوات في المثال السابق حيث ان القيم المبوبة ليست ارقام ذات اوزان مختلفة تدل عليها تكرارات القيم . وتكون الخطوة الاولى في العمل هي ايجاد مراكز الفئات ثم ايجاد لوغاريتمات هذه المراكز ثم ضرب لوغاريتم كل مركز \times تكرار المركز وجمع ناتج الضرب . وبقسمة هذا المجموع على مجموع التكرارات نحصل على لوغاريتم الوسط الهندسي ومن جدول الاعداد المقابلة نحصل على القيمة المقابلة لهذا اللوغاريتم تكون هي الوسط الهندسي .

مثال ١٦ :

ف	ك	م	لو م	لوم × ك
صفر-	٢٦٢	٥	٦٩٨٩٧	١٨٣١٣٠١٤
-١٠	٣٣٣	١٥	١١٧٦٠٩	٣٩١١٦٣٧٩٧
-٢٠	٢١٠	٢٥	١٣٩٧٩٤	٣٦٣٥٦٧٤٠
-٣٠	١٤٢	٣٥	١٥٤٤٠٧	٢١٩٢٥٧٩٤
-٤٠	٨٥	٤٥	١٦٥٣٢١	١٤٠٥٢٣٨٥
-٥٠	٣٢	٥٥	١٧٤٠٣٦	٥٥٦٩١٥٢
-٦٠	٢٠	٦٥	١٨١٢٩١	٣٦٢٥٨٢٠
-٧٠	١٥	٧٥	١٨٧٥٠٦	٢٨١٢٥٩٠
-٨٠	١٢	٨٥	١٩٢٩٤٢	٢٣١٥٣٠٤
-٩٠	٨	٩٥	١٩٧٧٧٢	١٥٨٢١٧٦
-١٠٠	٢	١٠٥	١٠٢١١٩	٤٠٤٢٤٨
	١١٢١			١٣٩١٢٠٩١٠

$$\therefore \text{لو الوسط الهندسي} = \frac{١٣٩١٢٠٩١}{١١٢١} = ١٢٤١٠٤$$

$$\therefore \text{الوسط الهندسي} = ١٧٤١٨$$

الوسط التوافقي

الوسط التوافقي لأي عدد من القيم هو مقلوب الوسط الحسابي لمقلوب القيم ، فإذا كانت القيم هي ١ س ، ٢ س ، ٣ س ، من

$$\therefore \text{الوسط التوافقي} = \frac{ن}{\frac{١}{١س} + \frac{١}{٢س} + \frac{١}{٣س} + \frac{١}{\dots\dots\dots} + \frac{١}{نس}}$$

$$\frac{1}{\text{م.ح}} = \frac{1}{\text{ن}} = \frac{\frac{1}{\text{م.ح}}}{\text{ن}} = \frac{1}{\text{الوسط التوافقي}} \quad \text{وبشكل آخر يكون}$$

مثال ١٧ :

$$\frac{3}{7} = \frac{3}{\frac{1}{8} + \frac{1}{4} + \frac{1}{2}} = 8,4,2 \quad \text{الوسط التوافقي للقيم}$$

$$3 \times 8,4,2 =$$

واذا كانت القيم مبوبة في توزيع تكراري نجد أولاً مراكز الفئات ثم نحسب مقلوباتها من الجدول الخاص بذلك ، ثم نضرب كل مقلوب في تكرار الفئة ثم نجمع حواصل الضرب ثم نقسم مجموع التكرارات على الناتج تكون النتيجة هي الوسط التوافقي كما هو مبين في المثال التالي :

مثال ١٨ :

ف	ك	م	مقlobات المراكز	المقlob × التكرار
صفر—	٢٦٢	٥	٠,٢	٥٢,٤
١٠—	٣٣٣	١٥	٠,٠٦٦٦	٢٢,١٧٧٨
٢٠—	٢١٠	٢٥	٠,٠٤	٨,٤
٣٠—	١٤٢	٣٥	٠,٠٢٨٦	٤,٠٦١٢
٤٠—	٨٥	٤٥	٠,٠٢٢٢	١,٨٨٧٠
٥٠—	٣٢	٥٥	٠,٠١٨٢	٠,٥٨٢٤
٦٠—	٢٠	٦٥	٠,٠١٥٤	٠,٣٠٨٠
٧٠—	١٥	٧٥	٠,٠١٣٣	٠,١٩٩٥
٨٠—	١٢	٨٥	٠,٠١١٨	٠,١٤١٦
٩٠—	٨	٩٥	٠,٠١٠٥	٠,٠٨٤٠
١٠٠—٢١١٠		١٠٥	٠,٠٠٩٥	٠,٠١٩٠
	١١٢١			٩٠,٢٦٠٥

$$\therefore \text{الوسط التوافقي} = \frac{1121}{90,2605} = 12,419$$

ويمكن التعبير عن الخطوات السابقة بالقانون :

$$\text{الوسط التوافقي} = \frac{\text{ن أو مح ك}}{\frac{\text{ك}}{\text{مح س}}}$$

ترجيح المتوسطات :

إذا طلب منا ان نحسب متوسط سعر القمح على اختلاف أنواعه وإذا علمنا ان سعر الكيلو من أصنافه الثلاث هي ٤٠ ، ٤٣ ، ٤٥ قرشاً فإنه من السهل ان نحسب المتوسط المطلوب يجمع هذه الأسعار وقسمة المجموع على عدد الأصناف ويسمى المتوسط في هذه الحالة بسيطاً حيث أننا لم نفرق بين أهمية صنف وآخر وهذا ما يجعل المتوسط الناتج لا يعبر تعبيراً دقيقاً عن متوسط سعر القمح في السوق نظراً لاختلاف أصنافه المختلفة ، ذلك الاختلاف الذي يظهر في كمية المبيع من هذه الأصناف. فإذا فرض ان نسبة ما يبيع من الصنف الاول الى ما يبيع من الصنف الثاني الى ما يبيع من الصنف الثالث هي كنسبة ٣ : ١ : ٦ فمتوسط السعر حينئذ لا بد أن يرجح وفقاً لتلك الأهمية هكذا:

$$= \frac{6 \times 45 + 1 \times 43 + 3 \times 40}{10} = 43,3 \text{ قرشاً .}$$

والمتوسط بهذه الطريقة لا بد ان يختلف عن المتوسط السابق ولكنه يكون أكثر دقة منه .

ومن الواضح ان كل متوسط يحسب لتوزيع تكراري هو في الواقع متوسط مرجح وأساس الترجيح وبمعنى آخر الأوزان المستخدمة في الترجيح

هي تكرارات الفئات المختلفة ، ذلك لأننا بضرب مركز الفئة \times تكرارها
كأننا نرجع هذا المركز وهو القيمة بعدد تكرارات هذه القيمة .

ولهذا لو اقتضى الأمر ترجيح الوسط حسابي أو الهندسي أو التوافقي
لعدة قيم بسيطة فيتبع لحساب هذه المتوسطات الطريقة التي شرحتها عند
الكلام على حساب متوسطات التوزيعات التكرارية .

ومن أهم ميادين استخدام فكرة الترجيح - حساب متوسط ما طرأ على
نفقات المعيشة من اختلاف بين تاريخ وآخر فلو علمنا أن أسعار الأغذية
ارتفعت في سنة ١٩٤٩ عنها في سنة ١٩٣٩ بمقدار ٣٠٠٪ وان إيجارات
المساكن زادت بنسبة ٢٠٠٪ وأسعار الملابس بمقدار ١٥٠٪ عما كانت عليه
فلا بد لكي نحسب متوسط ما طرأ على نفقات تلك البنود الثلاثة من تحديد
الأهمية النسبية لها ضمن نفقات الشخص العادي وتستخدم تلك الأهمية النسبية
أساساً للترجيح (يمكن معرفة الأهمية النسبية لنفقات الشخص العادي على
أبواب الأنفاق المختلفة من بحث لميزانية الأسرة يقوم على أساس المعاينة) .

وبما يستخدم فيه مبدأ الترجيح أيضاً - حساب معدلات الوفيات - فلو
علمنا أن معدل وفيات الأطفال في بلد ما هو ٧٠٪ ومعدل وفيات الشباب
١٩٪ والكحول ٥٠٪ فإنه يستحيل علينا حساب متوسط معدل الوفاة العام دون
تحديد الأهمية العددية النسبية للأطفال والشباب والشيوخ في المجتمع موضوع
الدراسة لاستخدامها في ترجيح تلك المعدلات .

خصائص المتوسطات المختلفة ومجالات استخدامها ،

الوسط الحسابي - يستخدم الوسط الحسابي في أغلب الحالات نظراً
لبساطة ووضوح فكرته الأساسية ولشيوع استعماله . وهو يمتاز عن المتوسطات
الأخرى من الناحية الفنية لامكان معالجته رياضياً ، الأمر الذي يجعل منه
وسيلة قوية في البحث الإحصائي .

ومن أهم تلك الصفات كون المجموع الجبري لانحرافات القيم عنه يساوي صفراً وهذه الخاصية هي الأساس الذي تبني عليه الطريقة المختصرة في حساب هذا المتوسط .

كذلك نلاحظ أن مجموع مربع الانحرافات القيم عن الوسط الحسابي أقل ما يمكن أي أقل من مجموع مربع الانحرافات القيم عن أي قيمة أخرى غير المتوسط . وسوف تساعدنا هذه الخاصية في قياس التشتت بين القيم .

كذلك نلاحظ أن الوسط الحسابي لمجموعة مكونة من مجموعات صغيرة يساوي الوسط الحسابي المرجح لمتوسطات تلك المجموعات الجزئية وهذه صفة لا تتوفر في غيره من المتوسطات .

كما أن الوسط الحسابي لمجموعة من القيم إذا ضرب في عدد هذه القيم فإن الناتج يساوي مجموع هذه القيم .

ومن أهم عيوب الوسط الحسابي كونه يتأثر بجميع القيم في المجموعه على حد سواء فإذا تصادف وجود قيم شاذة في المجموعة انحراف الوسط الحسابي تبعاً لذلك . ولما كانت القيم الشاذة في أي توزيع تكراري توجد غالباً في الطرفين لهذا يمكننا أن نقول أن الوسط الحسابي يتأثر بالقيم المتطرفة ولا يجب استخدامه في مثل هذه الحالات .

ولا يمكن تقدير الوسط الحسابي لتوزيع تكراري إذا كانت إحدى فئاته مفتوحة إذ يلزم العمل الحسابي لتقديره تعيين مراكز الفئات جميعها ولا يمكننا بطبيعة الحال تحديد مركز فئة لم يعين لها أحد حديها (إلا في الحالات التي تسمح طبيعة البيانات افتراض حد مقبول لها) .

ومما يجب الإشارة إليه أنه تبعاً لإحدى صفات الوسط الحسابي يمكن حسابه دون ضرورة معرفة القيم كلها واحدة واحدة إذ يكفي أن نعرف مجموع هذه القيم وعددها لكي نقدر وسطها الحسابي . فإذا أمكن الحصول على

مجموع ما يدفع لعمال مصنع معين من الأجور وعدد هؤلاء العمال امكن بقسمة الأول على الثاني الحصول على متوسط الاجر في هذا المصنع وهو ما لا يمكن عمله فيما يختص بالمتوسطات الأخرى .

الوسيط - أهم ما يمتاز به الوسيط عن غيره من المتوسطات امكان تقديره بدقة للتوزيعات التكرارية ذات الفئات المفتوحة وكذلك للتوزيعات المختلفة المدى أن التوزيعات الغير منتظمة - وكذلك يمكن حسابه بالرسم البياني بدقة.

كما وان الوسيط يمتاز بكونه لا يتأثر بالقيم الشاذة في التوزيعات فهو دائماً يقع بين الوسط الحسابي والمنوال في التوزيعات المعتدلة الالتواء . وهي خاصية سوف نشرحها بالتفصيل عند الكلام على العلاقة بين المتوسطات الثلاث - الوسط الحسابي والمنوال والوسيط .

ونلاحظ أن الوسيط له خاصية هامة وهي أن مجموع الانحرافات المطلقة للاعداد عن وسيطها أصغر من مجموع انحرافات المطلقه عن أي عدد آخر (الصناعة التي عند الوسيط الجغرافي) .

ويستعمل الوسيط في التعبير عن متوسط التراتيب حيث يكون من غير المتطوق استخدام الوسط الحسابي لأنه يتفق أكثر مع الاعداد التي تعبر عن القياس والم . كذلك يستخدم الوسيط في حالة التجارب التي تسبب ائتلاف المفردات المختبره أثناء اجراء التجربة مثل تجربة خاصة بمتوسط عمر اللبنة الكهربائية .

المنوال - يمتاز المنوال ببساطة فكرته ووضوحها اذ يمكن فهم معناه كمتوسط دون كبير عناء وقد يكون في هذه الصفة اسبق جميع المتوسطات الأخرى حتى الوسط الحسابي، فمن السهل أن ندرك معنى القيمة الأكثر تكراراً في مجموعة من القيم .

وكذلك يتميز المنوال بأنه لا يتأثر بالقيم الشاذة المتطرفة لأنه غالباً يكون وسط التوزيع فلا يتأثر بالأطراف .

على انه يجب أن نلاحظ ان المنوال لا يكون له أي دلالة إذا لم يكن هناك اتجاه واضح نحو التركيز في التوزيع التكراري . وكذلك نلاحظ أنه قيمة تقريبية وقد سبق الإشارة إلى هذه الصفة للمنوال ، ولهذا لا يمكن استخدامه في المسائل التي تقتضي معالجة جبرية ، ولذا فانه يحتفي في الموضوعات الاحصائية الراقية بعكس الحال في الوسط الحسابي والوسط الهندسي .

ويستحسن عدم استخدام المنوال في الحالات التي يكون التوزيع فيها مفتوحاً فقيمه تعتمد على التكرارات وهذه تتوقف على أطوال الفئات فلا بد من الوقوف على طول الفئة قبل الحكم على كونها منوالية أم لا . ولذلك يستحسن عدم استخدام المنوال في التوزيعات الغير متساوية الفئات حيث أن الطرق التي مبرحناها سابقاً تفترض ضمناً التساوي في الفئات .

ويعاب على المنوال انه يتأثر بطريقة اختيار فئات التوزيع أي ان القيمة التي نعتبرها ممثلة للمجموعة تتوقف على طريقة تلخيص الباحث لبياناته وهذه حقيقة تجعل المنوال لا يوثق به كممثل للمجموعة .

ولا يصح استخدام المنوال في حالة التوزيعات التكرارية الملثوية وخصوصاً الحادة الالتواء منها فكما أن الوسط الحسابي ينحرف كثيراً في اتجاه ذيل التوزيع نجد أن المنوال ينحرف الى الجهة الأخرى .

الوسط الهندسي - يخضع الوسط الهندسي للمعالجة الجبرية مثل الوسط الحسابي ، فان حاصل ضرب مجموعة من القيم لا يتغير إذا استبدل كل من هذه القيم بالوسط الهندسي ، فحاصل ضرب القيم ٢ ، ٤ ، ٨ ، ٦٤ والوسط الهندسي لهذه القيم يساوي ٤ ، فاذا ضربنا ٤ في نفسها ثلاث مرات فان الناتج يكون ٦٤ .

الحج وهو الوسط الهندسي لقيم الصغرى والكبرى
والقيم الصغرى متصفاً للوسط

وكما انه في حالة الوسط الحسابي يتساوى مجموع انحرافات القيم التي تقل عن الوسط الحسابي مع مجموع انحرافات القيم التي تزيد عنه ، فان حاصل ضرب نسب القيم التي تقل عن الوسط الهندسي إلى هذا الوسط تساوي حاصل ضرب نسبة القيم التي تزيد عنه . فمثلا القيم ٣ ، ٦ ، ٨ ، ٩ الوسط الهندسي لها = ٦ وبذلك يكون $\frac{3}{6} \times \frac{8}{6} = \frac{6}{6} \times \frac{6}{6}$.

وتبعاً لهذه الخاصية يستخدم الوسط الهندسي في حساب متوسطات النسب (الأرقام القياسية على وجه الأخص) وإذا استخدم الوسط الحسابي بدلاً من الهندسي لايجاد متوسط النسب فان النتيجة ستكون أعلى من الحقيقة ولهذا يقال ان الوسط الحسابي متحيز إلى أعلى ، فلو فرض مثلاً ان سعر سلعة ما قد زاد إلى الضعف ونقص سعر سلعة اخرى الى النصف فمن المعقول أن نقرر أن مستوى سعر السلعتين معاً لم يتغير ولكن لو استخدمنا الوسط الحسابي لظهر ان مستوى السعر لهاتين السلعتين قد زاد إلى $1\frac{1}{4}$ مرة عما كان عليه

المستوى القديم $\frac{1}{2} + \frac{1}{2} = \frac{1}{1} = 1$ في حين أن الوسط الهندسي لمستوى السعيرين هو نفسه $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

المستوى القديم $\sqrt{\frac{1}{2} \times \frac{1}{2}} = 1$. كذلك يستخدم الوسط الهندسي في

الموضوعات التي تستلزم حساب متوسط قيم متزايدة أو متناقصة . فمثلاً يمكن حساب معدل الزيادة في مبلغ معين من المال اودع بفائدة مركبة وذلك باستخدام المعادلة .

المبلغ في نهاية المدة = المبلغ في أول المدة (١ + r)ⁿ

r هي معدل الفائدة

n هي المدة .

١٠ هي لمدة

مثال ١٩ :

إذا أودع مبلغ ١٠٠٠ جنيه بفائدة مركبة في بنك لأصبح ١٦٠٠ جنيه بعد ١٢ سنة فكم يكون معدل الفائدة .

يظهر من ذلك أن هناك زيادة بمعدل ٦٠٪ في ١٢ سنة أي ٥٪ في السنة ، هذا إذا استخدمنا الوسط الحسابي ؛ ولكن الواقع أن هذا ليس معدل الفائدة الحقيقية ، أما إذا استخدمنا فكرة الوسط الهندسي يكون معدل الفائدة كالتالي :

$$1600 = 1000 (1+r)^{12}$$

$$\frac{1600}{1000} = (1+r)^{12}$$

$$1.6 = (1+r)^{12}$$

$$1.6^{1/12} = 1+r$$

$$1.04 = 1+r$$

$$r = 0.04$$

فإذا أردنا أن نحس قيمة هذا المبلغ بعد ستة سنوات بنفس الفائدة المركبة نجد أنه يساوي ١٠٠٠ (١ + ٠٠٤ ر) = ١٢٦٥ جنيه - هذه النتيجة يمكن الحصول عليها بإيجاد الوسط الهندسي للمبلغين ١٠٠٠ ، ١٦٠٠ فيكون :

$$1265 = 1000 \times 1600 \sqrt{}$$

مثال ٢٠ :

إذا طلب معرف مبلغ أو عدد وعرفنا العدد الذي أول المدة وفي آخره المطلوب في النصف فهو بالوسط الهندسي

إذا فرضنا أن عدد السكان في سنة ١٩٢٧ = ٩ مليون وعدد السكان في سنة ١٩٣٧ = ١٦ مليون

فإذا أردنا إيجاد عدد السكان في سنة ١٩٣٢ فهناك طريقتان :

الطريقة الأولى : عدد السكان في ١٩٣٢ = $9 \times 16 \sqrt{}$ = ١٢ مليون

الطريقة الثانية : $9 = 16 (1 + r)$

$$\text{لو } 16 = 9 \text{ لو } 10 + 1 \text{ لو } (r + 1)$$

$$\underline{17204120} = 900343 + 10 \text{ لو } (r + 1)$$

$$10 = 0.248777 \text{ لو } (r + 1)$$

$$\underline{170591} = r + 1$$

$$\underline{170591} = r$$

$$\underline{170591}$$

$$\underline{170591}$$

وبذلك يكون معدل زيادة السكان = 59.1 في الألف .

$$9 = 1932 (1 + 0.0591)$$

$$= 12 \text{ مليون .}$$

(استخدمنا فكرة الوسط الهندسي لأن الزيادة في عدد السكان تشبه تماماً الزيادة في مبلغ من المال بفائدة مركبة) .

وأمم ما يعاب على الوسط الهندسي صعوبته الحسابية وضرورة استخدام جداول اللوغاريتمات في ايجاده .

الوسط التوافقي - يقتصر استخدامه في ايجاد متوسطات الأسعار اذا اعطيت بدلالة وحدة النقود ، كذلك في حالة ايجاد متوسط السرعة اذا انها أيضاً تعطي في العادة بدلالة وحدة الزمن فيقال مثلاً (خمسة وحدات من سلعة ما بقرش - أو قطعت سيارة المسافة بسرعة 90 كيلومتراً في الساعة وهكذا) . لو استخدم الوسط الحسابي في أمثال تلك الحالات فانه يكون متحيزاً الى أعلى ، فلو فرضنا اننا اشترينا عدداً من الليمون بسعر ثلاثة بقرش ونفس العدد بسعر خمسة بقرش فمن المعقول ان نقرر ن متوسط سعر ما اشتريناه هو أربعة بقرش . نحصل على هذه النتيجة اذا استخدمنا الوسط التوافقي ولكن نحصل على سعر أعلى لو استخدمنا الوسط الحسابي هكذا :

$$\frac{1}{\frac{1}{3} + \frac{1}{3}} = 2 \div \left(\frac{1}{3} + \frac{1}{3} \right) = \frac{1}{\frac{2}{3}} = \frac{3}{2} = 1.5$$

وهذا العدد أكبر من $\frac{1}{4}$ وهو الجواب الصحيح . والوسط التوافقي هو

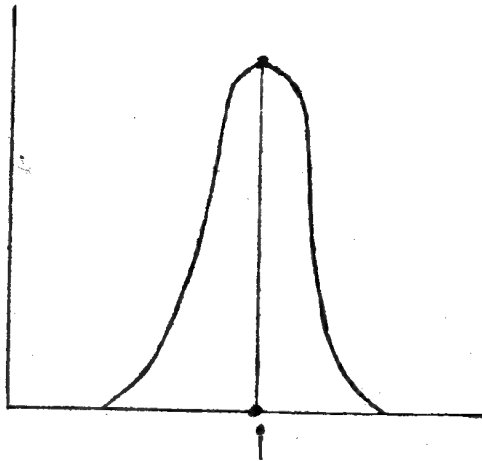
$$\frac{1}{\frac{1}{4}} = \frac{2}{5 + 3}$$

وعيب الوسط التوافقي عدم شيوع استخدامه وغرابته وصعوبة عملياته الحسابية اذ يقتضي استخدام جداول مقلوبات الأعداد وهذه قل ان تتوفر للباحث العادي .

العلاقة بين المتوسطات :

تتوقف العلاقة بين المتوسطات الثلاث - الوسط الحسابي والوسيط والمنوال - على نوع التوزيع التكراري وبالتالي على المنحنى الذي يمثل هذا التوزيع .

فإذا كان التوزيع تام التماثل يمثله منحنى معتدل طبيعي فإن المتوسطات الثلاث السابقة تتساوى تماماً أي تنطبق على بعضها عند القيمة المقابلة لقمة المنحنى التكراري الذي يمثل التوزيع كما يظهر في الشكل الآتي .



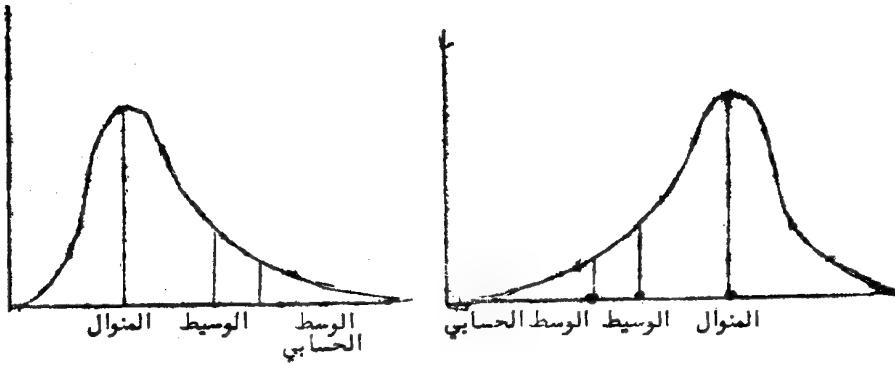
(النقطة أ تمثل الوسط الحسابي والوسيط والمنوال)

الوسط الحسابي - المنوال = ٣ (الوسط - الوسيط)

وإذا كان التوزيع معتدل الالتواء فن الوسط يقع دائماً بين المتوسطين الآخرين والمسافة بينه وبين الوسط الحسابي تساوي ثلث المسافة بين الوسط الحسابي والمنوال ، وهذه العلاقة يمكن وضعها في صيغة المعادلة الآتية :

$$\text{الوسط الحسابي} - \text{المنوال} = ٣ (\text{الوسط} - \text{الوسيط})$$

وتتضح هذه العلامة من الرسم البياني الآتي :



تمارين

١ - إذا كان متوسطات الدرجات في إحدى المواد لثلاث فصول هي ٧٥ ، ٨٢ ، ٨٤ وكان عدد التلاميذ في هذه الفصول ٣٢ ، ٢٥ ، ١٧ - احسب المتوسط العام لدرجات الثلاث فصول : الوسط = ٧٨ .

٢ - إذا كان المتوسط العام لاجور العمال = ٥٠٠٠ ليرة في العام وكان متوسط الاجور للعمال الذكور = ٥٢٠٠ ليرة ومتوسط الاجور للعمال الاناث = ٤٢٠٠ ليرة احسب نسبة العمال الذكور والاناث - النسبة هي ٨٠٪ ، ٢٠٪ .

٣ - التوزيع الآتي يبين قوة احتمالات الكابلات التي أنتجتها إحدى المؤسسات .

عدد الكابلات	قوة الاحتمال (طن)
٢	٩ر٣ - ٩ر٧
٥ر	١٠ر٢ - ٩ر٨
١٢	١٠ر٧ - ١٠ر٢
١٧	١١ر٢ - ١٠ر٨
١٤	١١ر٧ - ١١ر٣
٦	١٢ر٢ - ١١ر٨
٣	١٢ر٧ - ١٢ر٣
١	١٣ر٢ - ١٢ر٨

احسب متوسط قوة احتمال الكابل الواحد بالطريقة المطولة والطريقة المختصرة :

$$\text{س-} = ١١ر٠٩ \text{ طن}$$

٤ - الآتي التوزيع النسبي للمبات الكهربائية حسب طول عمرها :

عمر اللعبة بالساعة	% من مجموع اللبات
١٠٠ -	٥ر٦
٢٠٠ -	١٤ر٤
٣٠٠ -	١٤ر٦
٤٠٠ -	٢٥ر٤
٥٠٠ -	١٩ر٤
٦٠٠ -	١٦ر٦
٧٠٠ - ٨٠٠	٤ر٠

احسب الوسط الحسابي والوسط الهندسي والوسط التوافقي .

٥ - الآتي توزيع أعمار أرباب الاسر :

عدد الاسر بالآلف	العمر بالسنوات الكاملة
٢	أقل من ٢٥
٤	٢٥ - ٢٩
٥	٣٠ - ٣٤
١٠	٣٥ - ٤٤
٩	٤٥ - ٥٤
٦	٥٥ - ٦٤
٤	٦٥ - ٧٤
١	٧٥ وأكثر

احسب الوسيط والربيعين (حسابياً وبالرسم البياني)

٦ - احسب المتوال للتوزيع الآتي (بالطرق المختلفة)

التكرار	الفئات
٣	١٠ -
٧	١٥ -
١٦	٢٠ -
١٢	٢٥ -
٩	٣٠ -
٥	٣٥ -
٢	٤٠ - ٤٥

المتوال = ٢٣٥

X

المقاييس الثلاثة هي :

الفصل السادس

التشتت

- ① مقياس لقياس لنا مدى تباين أو تركز القيم حول القيمة المتوسطة أي مدى لفرقة القيم أو اختلافها أو تشتتها
- ② درجة تماثل القيم حول القيمة المتوسطة
- ③ درجة تفرطح التوزيع أو بمعنى

كان اهتمامنا في الفصول السابقة موجهاً الى كيفية تلخيص عدد كبير من البيانات الرقمية تلخيصاً يساعد على فهمها وتحليلها احصائياً ، ثم الى محاولة البحث عن قيمة واحدة يمكن ان توصف بها هذه البيانات وهي تلك القيمة التي نعتقد ان القيم الاخرى تميل الى التركيز حولها . ولكن أي متوسط لا يكفي وحده لقياس هذا الاتجاه نحو التركيز ، ولذلك نذكر في حاجة الى ثلاث أنواع أخرى من المقاييس حتى نستطيع ان نصف التوزيع التكراري وصفاً كاملاً وحتى يمكن مقارنته بتوزيع تكراري آخر . هذه المقاييس الثلاث هي أولاً مقياس يقيس لنا مدى تباعد أو تركز القيم حول القيمة المتوسطة وبمعنى آخر مدى بعثرة القيم أو اختلافها أو تشتتها ، والمقياس الثاني يقيس لنا درجة تماثل القيم حول القيمة المتوسطة والمقياس الثالث يقيس لنا درجة تفرطح التوزيع وبمعنى آخر درجة تجمع القيم عند القيمة المنوالية .

أهمية قياس التشتت

التشتت صفة هامة من صفات أي مجموعة من البيانات الرقمية فلا يمكن أن نتصور تساوي أطوال جميع الطلبة أو أوزانهم أو أعمارهم ولا يمكن ان نتصور تساوي الانتاج في جميع المؤسسات الصناعية أو تساوي دخول جميع أفراد المجتمع ، ولهذا فان القيمة التي نعتبرها ممثلة لمجموعة من القيم لا بد وان

تكون مصحوبة بقيمة أخرى تقيس لنا مدى تباعد القيم أو قربها من هذه القيمة ، بحيث اذا كبر هذا المقياس الى درجة بعيدة فان المتوسط يفقد أهميته كقيمة ممثلة ، أما اذا صغر مقياس التشتت تزداد أهمية المتوسط كقيمة ممثلة ، فيمكننا الاعتماد عليه في بحثنا لهذه البيانات بحثاً احصائياً

مقاييس التشتت

يمكن قياس التشتت بمقاييس مميزة بنفس وحدات البيانات التي نقوم بدراستها أو بمقاييس مطلقة أي كنسبة مئوية مستقلة عن الوحدة التي تقاس بها البيانات . فاذا استعملنا مقياساً من النوع الأول فاننا بذلك نقيس التشتت المطلق ، واذا استعملنا مقياساً من النوع الثاني فاننا نقيس التشتت النسبي . وبالطبع تتميز مقاييس النوع الثاني بأنها صالحة للمقارنة بين مجمرعات مختلفة بينما لا يمكن استخدام مقاييس النوع الأول في المقارنة .

وأبسط مقاييس التشتت هو المدى أي مدى تغير الظاهرة موضوع الدراسة . فلو أعطينا أجر مائة عامل مثلاً فأبسط مقياس للتشتت بين أجور هؤلاء العمال هو الفرق بين أكبر وأصغر أجر فكما زاد المدى أو الفرق بين هاتين القيمتين زاد عدم التجانس أي زاد تشتت المجموعة والعكس كلما نقص المدى .

وعيب هذا المقياس عدم دقته حيث انه يتوقف على قيمتين فقط فلو تصادف ان كانت قيم شاذة أصبح المدى لا يعبر تعبيراً دقيقاً عن التشتت بين البيانات . ولهذا فبالرغم من بساطة هذا المقياس الا أننا لا نستطيع أن نعتمد عليه في قياس التشتت أو في مقارنته بين مجموعات مختلفة وهذا هو السبب في أننا نلجأ الى مقاييس أخرى للتعبير عن التشتت لتتخلص بواسطتها من أثر القيم المتطرفة التي تكون أحياناً واضحة الشذوذ .

المدى الربيعي : الربيع الأدنى - الربيع الأعلى
 نصف المدى الربيعي : $\frac{1}{2} (\text{الربيع الأدنى} - \text{الربيع الأعلى})$

نصف المدى الربيعي

يعالج عيب المدى الذي أشرنا إليه بقياس التشتت بقياس لا يتأثر بالقيم المتطرفة وذلك بأخذ الفرق بين الربع الأعلى والأدنى . وحيث أننا نعلم أن نصف عدد القيم ينحصر بين الربعين وأن في هذا النصف يقع المتوسط الحسابي والوسيط والمنوال وهي القيم التي اعتبرناها ممثلة للتوزيع ، لهذا يكون من المنطق أن نعتمد إلى حد ما على قياس التشتت بواسطة نصف الفرق بين الربع الأعلى والأدنى .

ولحساب هذا المقياس نرتب القيم ترتيباً تصاعدياً أو تنازلياً ثم نحسب قيمة كل من الربع الأدنى والأعلى بالطريقة التي شرحناها في الفصل السابق ويكون :

المدى الربيعي = الربع الأعلى - الربع الأدنى .

$$\text{نصف المدى الربيعي} = \frac{1}{2} (\text{الربع الأعلى} - \text{الربع الأدنى}) .$$

التشتت حول المتوسط

في قياس التشتت بالطريقتين السابقتين كنا نقيس مدى الاختلاف بين القيم المختلفة التي تتخذها ظاهرة ما في تغيرها . ولكننا قد اتفقنا على أن تمثل هذه القيم بقيمة واحدة هي المتوسط ولذلك يكون من المنطق أن نحاول قياس تشتت القيم حول هذا المتوسط . وقد أشرنا في الفصل السابق إلى أن الوسط الحسابي يمكن أن يخضع للمعالجة الجبرية ولهذا نفضل قياس التشتت حوله .

الانحراف المتوسط :

هو متوسط مجموع الانحرافات القيم عن الوسط الحسابي ، ولكننا نعرف أن المجموع الجبري لانحرافات القيم حول الوسط الحسابي يساوي صفراً ، لهذا

الانحراف المتوسط : هو متوسط مجموع انحرافات القيم عن الوسط الحسابي .

يجب عند حساب الانحراف المتوسط أن نتخلص من الاشارات الجبرية بوسيلة ما ، وأبسط وسيلة لذلك هي مجرد اهمال هذه الاشارات .

وطريقة حساب الانحراف المتوسط في المجموعات الصغيرة (القيم البسيطة الغير مبوبة) هي إيجاد الوسط الحسابي للقيم ثم إيجاد انحرافات هذه القيم عن الوسط الحسابي ، ثم إيجاد مجموع هذه الانحرافات بعد اهمال اشاراتها ، وبقسمة هذا المجموع على عدد القيم ينتج الانحراف المتوسط .

فمثلاً في المجموعة ٣ ، ٦ ، ٩ ، ١٢ ، ١٥ ، ١٨ المتوسط = ٩

والانحرافات بدون اشارة هي ٦ ، ٣ ، ٠ ، ٣ ، ٦ ، ٩ فيكون الانحراف

$$\text{متوسط} = \frac{١٨}{٥} = ٣.٦$$

أما في حالة التوزيع التكراري فتكون خطوات العمل كالآتي اذا أردنا أن نحسب الانحرافات عن الوسط الحسابي رأساً .

١ - حساب الوسط الحسابي . ✓

٢ - حساب الانحرافات (اهمال الاشارة الجبرية) عن الوسط الحسابي . ✓

٣ - إيجاد حاصل ضرب تكرار كل فئة في انحراف مركز الفئة عن الوسط الحسابي بدون اشارة . ✓

٤ - جمع هذه الانحرافات وقسمة الناتج على التكرار الكلي فينتج لدينا الانحراف المتوسط .

وتتضح هذه الخطوات من الجدول الآتي :

مثال ٢١ :

الفئات	التكرار	المراكز	م × ك	الانحرافات	ح × ك
٠ - ٠	٢٠	٥	١٠٠ ✓	١٧	٣٤٠
- ١٠	٨٠	١٥	١٢٠٠	٧	٥٦٠
- ٢٠	٥٠	٢٥	١٢٥٠	٣	١٥٠
- ٣٠	٤٠	٣٥	١٤٠٠	١٣	٥٢٠
- ٤٠	١٠	٤٥	٤٥٠	٢٣	٢٣٠
	٢٠٠		٤٤٠٠		١٨٠٠

$$\text{الوسط الحسابي} = \frac{4400}{200} = 22$$

$$\text{الانحراف المتوسط} = \frac{1800}{200} = 9 \quad (\text{حسبت الانحرافات عن } 22)$$

ولكن عند حساب الانحرافات عن الوسط الحسابي قد توجد كسور كبيرة في بعض الحالات ، الأمر الذي يؤدي الى تعقيد العمليات الحسابية ، ولهذا نلجأ الى طريقة اخرى لتبسيط العمليات الحسابية لتفادي الكسور في الانحرافات وذلك باستخدام وسط فرضي وتطبيق علاقة جبرية خاصة :

مجموع الانحرافات عن الوسط الحسابي .

الوسط الحسابي - الوسط الفرضي .

مجموع التكرارات التي تقل عن الوسط الحسابي .

وبذلك تكون ح = ح + ف (ك - ك)

$$\text{حس} = \text{حج} + \text{ف} (ك - ك)$$

الأسئلة ٢ و ٣ (١٥ - ٢٥)

والجدول التالي يوضح كيفية استخدام هذه العلاقة في حساب الانحراف المتوسط :

مراكز الفئات	التكرار	ح عن وسط فرضي ٢٥	ح بدون اشارة	ح x ك
٩	٢٠	٢٠ -	٢٠	٤٠٠
١٥	٨٠	١٠ -	١٠	٨٠٠
٢٥	٥٠	صفر	صفر	صفر
٣٥	٤٠	١٠ +	١٠	٤٠٠
٤٥	١٠	٢٠ +	٢٠	٢٠٠
	٢٠٠			١٨٠٠

$$١٨٠٠ + (٣ -) = ١٨٠٠$$

$$٣ - = ٢٥ - ٢٢ = ٣$$

$$١٠٠ = ١$$

$$١٠٠ = ٣$$

$$١٨٠٠ = (٣ -) (١٠٠ - ١٠٠) + ١٨٠٠$$

$$١٨٠٠ =$$

$$\frac{١٨٠٠}{٢٠٠} = \text{الانحراف المتوسط} = ٩$$

وتفسير هذه الطريقة ان الرقم ١٨٠٠ هو عبارة عن مجموع انحرافات القيم عن وسط فرضي (٢٥) ، ولكننا نريد مجموع انحرافات القيم عن الوسط الحسابي ولهذا نكون في حاجة الى تصحيح الرقم ١٨٠٠ ، وللتصحيح نلاحظ اننا حسبنا انحرافات القيم التي تقل عن الوسط الحسابي (٢٢) عن وسط فرضي (٢٥) ولذلك لا بد من طرح $١٠٠ \times (٢٥ - ٢٢) = ٣٠٠$ أما القيم التي تزيد عن الوسط الحسابي فقد حسبنا انحرافاتنا عن الوسط الفرضي ٢٥ ولذلك يجب ان نضيف $١٠٠ \times (٢٥ - ٢٢) = ٣٠٠$ وبذلك يجب تصحيح مجموع الانحرافات عن الوسط الفرضي بطرح $١٠٠ \times ٣ -$ وجمع $١٠٠ \times ٣ +$.

المعياري

(Standard Deviation) الانحراف المعياري

وهو المقياس الثاني الذي يقيس تشتت القيم حول الوسط الحسابي. والفكرة الأساسية في هذا المقياس هو انه بدلا من إهمال الاشارات الجبرية عند حساب الانحراف المتوسط وهو اجراء غير منطقي من الناحية الجبرية ، لهذا نحاول التخلص من هذه الاشارات بطريقة أخرى أكثر صلاحية وذلك بتربيع الانحرافات فيتحول السالب منها والموجب الى قيم موجبة .

ونلاحظ ان الانحراف المعياري هو أقوى مقاييس التشتت حساسية وأكثرها شيوعاً فتكاد جميع وسائل التحليل الاحصائي تعتمد عليه . ويمكن تعريفه بأنه الجذر التربيعي لمتوسط مجموع مربعات انحرافات القيم عن الوسط الحسابي . فإذا أخذنا القيم ٣ ، ٧ ، ٨ منه فان انحرافاتهما عن الوسط الحسابي هي ٣ ، ١ ، ٢ على التوالي ومربعات تلك الانحرافات هي ٩ ، ١ ، ٤

ومتوسط مجموع هذه الانحرافات هو $\frac{4 + 1 + 9}{3} = 4.67$ وجذره

التربيعي ٢.١٦

ويكتفي في بعض الأحيان عند قياس التشتت بالقيمة السابقة قبل حساب الجذر التربيعي لها أي ٤.٦٧ وفي هذه الحالة يطلق على المقياس اسم التباين . وواضح ان التباين ما هو الا مربع الانحراف المعياري .

ولا يعاب على الانحراف المعياري مقياس للتشتت سوى ان تميزه من تمييز المتغير الأصلي ولهذا لا يمكن استخدامه في مقارنة التشتت بين مجموعتين لهما تمييز مختلف ، فلو كان لدينا ظاهرتين لمجموعة واحدة كالأجور وعدد أبناء طائفة من العمال وأردنا مقارنة تشتت الأجور بتشتت عدد الأبناء لما أمكننا ذلك باستخدام الانحراف المعياري اذ ان الانحراف المعياري للأجور سيكون ه غروش مثلا بينما الانحراف المعياري لعدد الأبناء سيكون فردين مثلا ولا يمكننا بطبيعة الحال مقارنة هذا بذلك .

الانحراف المعياري الذي يقيس تشتت القيم حول الوسط الحسابي. والفكرة الأساسية في هذا المقياس هو انه بدلا من إهمال الاشارات الجبرية عند حساب الانحراف المتوسط وهو اجراء غير منطقي من الناحية الجبرية ، لهذا نحاول التخلص من هذه الاشارات بطريقة أخرى أكثر صلاحية وذلك بتربيع الانحرافات فيتحول السالب منها والموجب الى قيم موجبة . ونلاحظ ان الانحراف المعياري هو أقوى مقاييس التشتت حساسية وأكثرها شيوعاً فتكاد جميع وسائل التحليل الاحصائي تعتمد عليه . ويمكن تعريفه بأنه الجذر التربيعي لمتوسط مجموع مربعات انحرافات القيم عن الوسط الحسابي . فإذا أخذنا القيم ٣ ، ٧ ، ٨ منه فان انحرافاتهما عن الوسط الحسابي هي ٣ ، ١ ، ٢ على التوالي ومربعات تلك الانحرافات هي ٩ ، ١ ، ٤ ومتوسط مجموع هذه الانحرافات هو $\frac{4 + 1 + 9}{3} = 4.67$ وجذره التربيعي ٢.١٦ ويكتفي في بعض الأحيان عند قياس التشتت بالقيمة السابقة قبل حساب الجذر التربيعي لها أي ٤.٦٧ وفي هذه الحالة يطلق على المقياس اسم التباين . وواضح ان التباين ما هو الا مربع الانحراف المعياري . ولا يعاب على الانحراف المعياري مقياس للتشتت سوى ان تميزه من تمييز المتغير الأصلي ولهذا لا يمكن استخدامه في مقارنة التشتت بين مجموعتين لهما تمييز مختلف ، فلو كان لدينا ظاهرتين لمجموعة واحدة كالأجور وعدد أبناء طائفة من العمال وأردنا مقارنة تشتت الأجور بتشتت عدد الأبناء لما أمكننا ذلك باستخدام الانحراف المعياري اذ ان الانحراف المعياري للأجور سيكون ه غروش مثلا بينما الانحراف المعياري لعدد الأبناء سيكون فردين مثلا ولا يمكننا بطبيعة الحال مقارنة هذا بذلك .

والعيب الثاني للانحراف المعياري تأثره بالوسط الحسابي للمجموعة ولهذا لا يمكن استخدامه في المقارنة حتى بين مجموعتين لهما نفس التمييز اذا اختلف وسطها الحسابي ، فلو أخذنا عدداً من طلبة المدارس الابتدائية ومتوسط أعمارهم ٩ سنوات وحسبنا الانحراف المعياري للأعمار وكان سنتين مثلاً ثم أخذنا مجموعة أخرى من طلبة الجامعة ومتوسط أعمارهم ٢٠ سنة مثلاً ووجدنا ان الانحراف المعياري لأعمارهم هو ٥ سنوات مثلاً فلا يمكننا القول بأن تشتت المجموعة الأولى أقل من تشتت المجموعة الثانية حيث ان هذين المقياسين يتأثران بقيمة الوسط الحسابي في كل حالة . ولهذا لمقارنة التشتت بين مجموعتين يستحسن عدم استعمال الانحراف المعياري . وقد اشرنا الى كيفية حساب الانحراف المعياري لمجموعة غير مبوبة من القيم ، أما بالنسبة لتوزيع تكراري يكون من المناسب استعمال وسط فرضي بدلاً من الوسط الحسابي الحقيقي وذلك لتسهيل العمليات الحسابية ثم استعمال معادلة تصحيح . وتكون خطوات العمل كما يلي :

- ١ - نحدد مراكز الفئات ونأخذ أحد هذه المراكز وسطاً فرضياً .
- ٢ - نحسب انحرافات المراكز الأخرى عن هذا الوسط الفرضي - سيكون بعضها سالباً وبعضها موجباً .
- ٣ - تضرب تكرار كل فئة في الانحراف المقابل وتجمع نتائج الضرب جمعاً جبرياً .
- ٤ - تضرب حاصل الضرب مرة أخرى في الانحراف فننتج مجموعة أعداد كلها موجبة وتجمع هذه الأعداد .
- ٥ - نستخرج الانحراف المعياري باستخدام القانون .

$$\text{الانحراف المعياري } (O) = \sqrt{\frac{\sum (\frac{f \cdot d^2}{n})}{n} - \left(\frac{\sum f \cdot d}{n}\right)^2}$$

$$\sqrt{\frac{277}{450} - \left(\frac{277}{450}\right)^2}$$

والجدول التالي يوضح عملياً طريقة حساب الانحراف المعياري :

مثال ٢٢ :

الفئات	التكرارات	المراكز	ح عن وسط فرضي	ح × ك	ح ^٢ × ك
			١٠		
صفر -	٢٠	٥	٢ -	٤٠ -	٨٠
١٠ -	٨٠	١٥	١ -	٨٠ -	٧٠
٢٠ -	٥٠	٢٥	صفر	صفر	صفر
٣٠ -	٤٠	٣٥	١ +	٤٠ +	٤٠
٤٠ -	١٠	٤٥	٢ +	٢٠ +	٤٠
	<u>٢٠٠</u>			<u>٦٠ -</u>	<u>٢٤٠</u>

$$\text{الانحراف المعياري} = \sqrt{10 \left(\frac{240}{200} - \left(\frac{60}{200} \right)^2 \right)}$$

$$= \sqrt{10 - 12} = 0.9$$

$$= \frac{1.055}{1.11} \sqrt{10} = 1.055$$

نلاحظ اننا اختصرنا الانحرافات على ١٠ لأن الفئات كلها متساوية المدى ،
والمسدى = ١٠ . ولسبب الاختصار نضرب نتيجة الجذر التربيعي في ١٠
لنرجع الأرقام الى أصلها . على ان هذا الاجراء لا يجب أن تتبعه إلا في حالة
التوزيع التكراري المنتظم أي التوزيع الذي يتساوى فيه مدى الفئات وقد
أشرنا الى هذه الملاحظة في ايجاد الوسط الحسابي .

على ان الجواب الذي نحصل عليه باتباع هذه الطريقة المختصرة التي تقوم
على أساس ايجاد الانحرافات عن وسط فرضي هو نفس الجواب الذي يمكن

ان نحصل عليه لو اتبعنا الطريقة القائمة على أساس إيجاد الانحرافات عن الوسط الحسابي . وليس اتباعنا لهذه الطريقة المختصرة إلا من قبيل تسهيل العمل الحسابي فقط . فالقاعدة العامة للانحراف المعياري هي الجذر التربيعي لمتوسط مجموع مربع انحرافات القيم عن الوسط الحسابي ، ولكن في حالة إيجاد الانحرافات عن الوسط الفرضي لا بد من تصحيح هذه الانحرافات .

مناقشة قوانين الانحراف المعياري

لاحظنا ان القاعدة الأساسية للانحراف المعياري هي :

$$(1) \quad \sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

نستطيع ان نصنع هذا القانون في شكل آخر بتحليل $\sum (x - \bar{x})^2$.
نفترض أن لدينا القيم $x_1, x_2, x_3, \dots, x_n$ ، ووسطها الحسابي \bar{x} =

∴ انحراف القيمة الاولى $x_1 - \bar{x}$ مربع هذا الانحراف $= x_1^2 - 2x_1\bar{x} + \bar{x}^2$

∴ انحراف القيمة الثانية $x_2 - \bar{x}$ ومربع هذا الانحراف $= x_2^2 - 2x_2\bar{x} + \bar{x}^2$

وانحراف القيمة الثالثة $x_3 - \bar{x}$ ومربع هذا الانحراف $= x_3^2 - 2x_3\bar{x} + \bar{x}^2$

والانحراف القيمة الرابعة = س٤ - س٣ - س٢ - س١
مربع هذا الانحراف = س٢ - س٣ - س٢ - س١ + س٢ - س١

∴ مجموع مربع الانحرافات القيم = س٢ - س٣ - س٢ - س١ + س٢ - س١

∴ مجموع مربع الانحرافات القيم = س٢ - س٣ - س٢ - س١ + س٢ - س١

∴ س (س - س) = س٢ - س٣ - س٢ - س١

$$(٢) \quad \sqrt{\frac{س٢ - س٣ - س٢ - س١}{ن}} = \sqrt{\frac{س٢ - س٣ - س٢ - س١}{ن}} = ع$$

وإذا أخذنا الانحرافات عن الوسط الفرضي ويكون القانون :

$$\sqrt{\frac{س٢ (س - و) - س٣ (س - و) - س٢ (س - و) - س١ (س - و)}{ن}} = ع$$

حيث ان س (س - و) هي مجموع مربع الانحرافات عن الوسط
فرضي ولذا يجب تصحيحها حتى تساوي مجموع مربع الانحرافات عن الوسط
لحسابي ولهذا أضفنا الجزء (س - و) . على أن هذا القانون هو مجرد
كل جبري آخر للقوانين السابقة ولا ثبات ذلك نفترض أن لدينا القيم :

س١ ، س٢ ، س٣ ، س٤ ووسطها الفرضي = و

∴ انحراف القيمة الاولى عن و = س١ - و

ومربع هذا الانحراف = س١ - و + س١ - و

وانحراف القيمة الثانية عن و = س_٢ - و
ومربع هذا الانحراف = س_٢^٢ - ٢ و س_٢ + و^٢

وانحراف القيمة الثالثة عن و = س_٣ - و
ومربع هذا الانحراف = س_٣^٢ - ٢ و س_٣ + و^٢

وانحراف القيمة الرابعة عن و = س_٤ - و
ومربع هذا الانحراف = س_٤^٢ - ٢ و س_٤ + و^٢

∴ مجموع مربع هذه الانحرافات = محس^٢ - ٢ و محس + ن و^٢ أي
مح (س - و)^٢ = محس^٢ - ٢ و ن س⁻ + ن و^٢

$$\therefore \frac{\text{مح} (س - و)^2}{ن} = \frac{\text{محس}^2}{ن} - \frac{٢ و ن س^-}{ن} + \frac{٢ و^2}{ن}$$

$$\therefore ع = \sqrt{\frac{\text{محس}^2}{ن} - \frac{٢ و ن س^-}{ن} + \frac{٢ و^2}{ن}}$$

$$\therefore ع = \sqrt{\frac{\text{محس}^2}{ن} - \frac{٢ و ن س^-}{ن}} \quad \text{وهو نفسه القانون الثاني.}$$

وبذلك يكون القانون الذي نستخدمه عند استخدام الوسط الفرضي هو:

$$(٣) \quad ع = \sqrt{\frac{\text{مح} (س - و)^2}{ن} - \frac{٢ (س - و)}{ن}}$$

هذا القانون هو نفسه الذي نستخدمه في إيجاد الانحراف المعياري للتوزيع التكراري حيث أن σ^2 (س - و) يعبر عنها في التوزيع التكراري بالرمز σ^2 (ح ك) هي مربع الانحراف عن الوسط الفرضي وضربناها في ك لان القيم متكررة ولذلك فأنحرافاتنا ومربع انحرافاتنا تتكرر معها). أما ن فترمز الى مجموع التكرارات في التوزيع التكراري. أما σ^2 (س - و) فهي مربع الفرق بين الوسط الحسابي والوسط الفرضي وحسب قانون الوسط الحسابي للتوزيع التكراري

$$\sigma^2 = \frac{\sum H K}{\sum K}$$

$$\therefore \sigma^2 = \frac{\sum H K}{\sum K}$$

$$\therefore (\sigma^2 - \sigma^2) = \left(\frac{\sum H K}{\sum K} \right)^2$$

وبذلك يكون قانون الانحراف المعياري للتوزيع التكراري هو :

$$\sigma = \sqrt{\frac{\sum H K^2}{\sum K} - \left(\frac{\sum H K}{\sum K} \right)^2}$$

ولا يشتمل هذا القانون على الضرب في طول الفئة حيث أننا أشرنا قبل ذلك الى أننا نضرب في طول الفئة عندما نختصر الانحرافات فقط ولذلك فالضرب في طول الفئة ليس شيئاً أساسياً في القانون .

مثال ٢٣ :

س	س - س	س - س - س	س ^٢	س - و	(س - و) ^٢
٣	- ٢		٩	- ١	١
٤	- ١		١٦	صفر	صفر
٨	+ ٣		٦٤	+ ٤	١٦
صفر			٨٩		١٧

و = ٤

∴ ع بالقانون الاول = $\sqrt{\frac{٢(س - س - س)}{ن}} = \sqrt{\frac{١٤}{٣}}$ ٢١٦ =

ع وبالقانون الثاني = $\sqrt{\frac{٢محس - س}{ن}} = \sqrt{\frac{٨٩}{٣}} = ٢٥ - ٢١٦ =$

و ع بالقانون الثالث =

$\sqrt{\frac{٢(س - و) - س}{ن}} = \sqrt{\frac{١٧}{٣}} = ١ - ٢(١) = ٢١٦ =$

أي ان القوانين الثلاث تعطينا نفس النتيجة وليس ذلك بمستغرب فهذه القوانين ليست إلا صوراً جبرية مختلفة مشتقة من القانون الاول .

معامل الاختلاف :

أشرنا فيما سبق إلى أن الانحراف المعياري بالرغم من دقته وكثرة استعماله

معامل الانحراف = $\frac{\text{الانحراف المعياري}}{\text{الوسيط}} \times 100$

الوسيط الحسابي

معامل الانحراف هو النسبة المئوية للانحراف المعياري ونسبة الوسط الحسابي

إلا أنه لا يفيدنا في حالة مقارنة تشتت مجموعتين احصائيتين لأسباب شرحناها فيما تقدم . ولكي نتمكن من المقارنة يجب الاستغناء عن الوحدات المميزة باستخدام اعداد مجردة خالية من التمييز ، ونحصل على مثل هذه الاعداد بقسمة عددين من نوع واحد أي مميزين بنفس الوحدات . وأكثر مقاييس المقارنة استخداماً هو معامل الاختلاف ونحصل عليه بقسمة الانحراف المعياري للتوزيع على الوسط الحسابي له وضرب الناتج في ١٠٠ . وبمعنى آخر ليس معامل الاختلاف إلا النسبة المئوية للانحراف المعياري منسوباً إلى الوسط الحسابي - وبالطبع كلما كبر معامل الاختلاف كلما دل ذلك على قوة التشتت بين مفردات التوزيع والعكس كلما صغر دل ذلك على ضعف التشتت

كلما كبر معامل الانحراف دل ذلك على قوة التشتت
وكذا صغر دل ذلك على ضعف التشتت

ولكن الانحراف المعياري بالرغم من صلاحيته وأهميته لا يمكن حسابه إلا في التوزيعات التي نستطيع تحديد وسطها الحسابي وذلك يستحيل في بعض التوزيعات (التوزيعات المفتوحة) ، ولهذا نستعاض عن الانحراف المعياري بنصف المدى الربيعي (يسمى أحياناً الانحراف الربيعي) . ولكن يعاب على هذا المقياس أنه يميز بنفس وحدات التوزيع ولذا لا يمكن استخدامه في المقارنة . ويمكن علاج هذا العيب كما فعلنا بالنسبة للانحراف المعياري بنسبته إلى الوسيط . وبذلك يكون معامل الاختلاف الربيعي =

معامل الانحراف الربيعي = $\frac{\text{نصف المدى الربيعي}}{\text{الوسيط}} \times 100$

$$\frac{\text{نصف المدى الربيعي}}{\text{الوسيط}} \times 100$$

ولما كان الوسيط في معظم التوزيعات القريبة من التماثل يساوي نصف

معامل الانحراف الربيعي = $\frac{\text{نصف المدى الربيعي}}{\text{الوسيط}} \times 100$

$$\begin{aligned} & \frac{\text{الربيع الأعلى} - \text{الربيع الأدنى}}{2} \\ & 100 \times \frac{\text{الربيع الأعلى} + \text{الربيع الأدنى}}{2} \\ & \text{الاختلاف الربيعي} = \end{aligned}$$

وبذلك يكون معامل

$$100 \times \frac{\text{الربيع الأعلى} - \text{الربيع الأدنى}}{\text{الربيع الأعلى} + \text{الربيع الأدنى}} = \text{الاختلاف الربيعي}$$

معنى التشتت في التوزيع التكراري :

نلاحظ أن التوزيعات التكرارية المعتدلة أو القريبة من الاعتدال (المثالية) لها خاصية تفيدنا كثيراً في الدراسات الاحصائية (العينات) .

ذلك أن المساحة المحدودة بالمنحنى التكراري المعتدل يمكن تقسيمها الى ثلاث قطاعات ، القطاع الأول يشمل ٦٨ر٢٧ ٪ من مجموع الوحدات (التكرارات) وقيمة كل وحدة منها لا تختلف عن الوسط الحسابي للتوزيع في الحد الأقصى الا بمقدار الانحراف المعياري سواء بالزيادة أو بالنقص (س- ع ، ع ، س+ ع) . أما القطاع الثاني فيشمل ٩٥ر٤٥ ٪ من مجموع الوحدات وقيمة كل وحدة منها لا تختلف عن الوسط الحسابي للتوزيع في الحد الأقصى الا بمقدار ضعف الانحراف المعياري سواء بالزيادة أو بالنقص (س- ع٢ ، ع٢ ، س+ ع٢) . والقطاع الثالث ويشمل ٩٩ر٧٣ ٪ أي تقريباً مجموع الوحدات التي لا يمكن أن تختلف قيمة الوحدة منها عن الوسط الحسابي الا بمقدار ثلاثة أمثال الانحراف المعياري سواء بالزيادة أو بالنقص (س- ع٣ ، ع٣ ، س+ ع٣) .

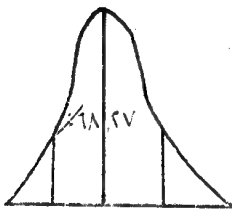
وبذلك نستطيع أن نقول أنه كلما نجد وحدة واحدة في التوزيع التكراري المعتدل تختلف قيمتها عن الوسط الحسابي للتوزيع الا بمقدار ع٣ سواء بالزيادة أو بالنقص .

الوسط الحسابي ٢٨٠
(س- ع) ٤
(س- ع٢) ٦
(س- ع٣) ٨
(س+ ع) ٤
(س+ ع٢) ٦
(س+ ع٣) ٨

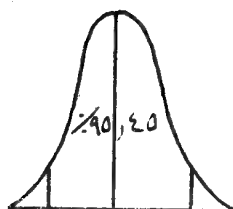
ويمكن توضيح الظاهرة السابقة بالجدول الآتي الذي يعطي مثلاً لتوزيع تكراري معتدل لأطوال مجموعة من الرجال بالبوصة .

الطول بالبوصة	عدد الرجال	س- ١ ع +	س- ٢ ع +	س- ٣ ع +
٦١ -	٢			
٦٢ -	٥			
٦٣ -	١٧			
٦٤ -	٤٣			
٦٥ -	٨٦			
٦٦ -	١٥٢			
٦٧ -	١٩٣			
٦٨ -	١٩٧			
٦٩ -	١٤٨			
٧٠ -	٩١			
٧١ -	٤٥			
٧٢ -	١٦			
٧٣ -	٤			
٧٤ - ٧٥	١			
المجموع	١٠٠٠			

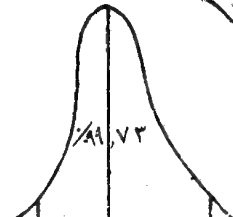
ويمكن توضيح ذلك بيانا كالآتي :



س- ١ ع + س- ٢ ع + س- ٣ ع +



س- ٢ ع + س- ٣ ع + س- ٤ ع +



س- ٣ ع + س- ٤ ع + س- ٥ ع +

وبمعنى آخر يمكن تقسيم مساحة المنحنى المعتدل (مجموع التكرارات في التوزيع المعتدل) الى قطاعات بدلالة الوسط الحسابي والانحراف المعياري لهذا التوزيع . فاذا كانت قيمة معينة في التوزيع تزيد عن الوسط الحسابي بمقدار ١ ع تكون المساحة المحصورة بين هذه القيمة وخط التماثل (الذي يمثل الوسط الحسابي للتوزيع) ٣٤ ٪ تقريباً من مجموع المساحة أي من مجموع التكرارات . واذا كانت القيمة تنقص عن الوسط الحسابي بمقدار ١ ع تكون المساحة المحصورة بين هذه القيمة وخط التماثل ٣٤ ٪ من مجموع المساحة ، أي أن المساحة المحصورة بين القيمتين تكون ٦٨ ٪ من مجموع المساحة تقريباً .

واذا كانت القيمة تزيد عن الوسط الحسابي بمقدار ٢ ع فإن المساحة المحصورة بين هذه القيمة وخط التماثل تكون ٤٧٥ ٪ تقريباً من مجموع المساحة . كذلك اذا كانت القيمة تقل عن الوسط الحسابي بمقدار ٢ ع تكون المساحة المحصورة بينها وبين خط التماثل ٤٧٥ ٪ أيضاً وبذلك يكون مجموع المساحة المحصورة بين القيمتين ٩٥ ٪ تقريباً من مجموع المساحة أي من مجموع التكرارات .

واذا كانت القيمة تزيد عن الوسط الحسابي بمقدار ٣ ع تكون المساحة المحصورة بينها وبين خط التماثل ٤٩,٨٦ ٪ تقريباً من مجموع المساحة . كذلك اذا كانت القيمة تقل عن الوسط الحسابي بمقدار ٣ ع تكون المساحة المحصورة بينها وبين خط التماثل ٤٩,٨٦ ٪ أيضاً . وبذلك يكون مجموع المساحة المحصورة بين القيمتين ٩٩,٧٣ ٪ من مجموع المساحة أي من مجموع التكرارات .

وحقن نحدد كم يكون الفرق بين القيمة والوسط الحسابي من وحدات ع

نستخدم القانون $\frac{s - s}{ع}$ (س رمز يدل على القيمة) ، تتوقف المساحة بين القيمة وخط التماثل على نتيجة هذا الكسر فإذا كانت ١ كانت المساحة ٣٤٪ وإذا كانت ٢ كانت ٤٧ر٥٪ وإذا كانت ٣ كانت المساحة ٤٩ر٨٦٪ من مجموع المساحة .

علاقات هامة بين مقاييس التشتت:

ونلاحظ انه بالنسبة للتوزيعات التكرارية القريبة من الاعتدال يكون الانحراف المتوسط $\frac{٤}{٥}$ الانحراف المعياري ، كما يكون نصف المدى الربيعي $\frac{٢}{٣}$ الانحراف المعياري . وتترتب هذه العلاقات على خاصية التوزيع المعتدل حيث يكون الانحراف المتوسط ٠٠ر٧٩٧٩ من الانحراف المعياري ، كما يكون نصف المدى الربيعي ٠ر٦٧٤٥ من الانحراف المعياري .

تمارين :

١ - احسب الانحراف المعياري للقيم : ١٢ ، ٦ ، ٧ ، ٣ ، ١٥ ، ١٠ ، ١٨ ، ٥ ، بثلاث طرق ع = ٤٨٧

كم يكون التباين لهذه القيم ؟

٢ - احسب الانحراف المعياري للتوزيع الآتي بطريقتين :

ف	ك
٦٢ - ٦٠	٥
٦٥ - ٦٣	١٨
٦٨ - ٦٦	٤٢
٧٤ - ٧٢	٨

ع = ٢٩٢

كم يكون التباين لهذا التوزيع ؟

٣ - احسب الانحراف المعياري للتوزيع الآتي :

معدلات الذكاء ١١٤ ١١٠ ١٠٦ ١٠٢ ٩٨ ٩٤ ٩٠ ٨٦ ٨٢ ٧٨ ٧٤ ٧٠

عدد التلاميذ ١٨ ٢٧ ٣٨ ٥٤ ٧٢ ٨٥ ٦٦ ٤٥ ٢٨ ١٦ ٩ ٣٠

احسب عدد التلاميذ الذين تتراوح معدلات ذكائهم بين س-١ + ع ١ .

» » » » » » » س-٢ + ع ٢ .



الفصل السابع

الالتواء

التواء أي توزيع يمثل انعدام التماثل فيه فجميع التوزيعات المتماثلة أو الطبيعية ينعدم فيها الالتواء نظراً لانتظامها أو تماثلها حول نقطة التركيز فيها ، فوجود الالتواء دليل انعدام الانتظام في التوزيع .

والالتواء أنواع فمن التوزيعات ما كان التواءه معتدلاً ومنها ما كان التواءه حاداً ومنها ما كان التواءه موجباً ومنها كان التواءه سالباً ، ولقد أشرنا الى جميع تلك الأنواع عند الكلام على أشكال المنحنيات التكرارية .

ويمكن الوقوف على طبيعة ودرجة التواء أي توزيع بمجرد النظر الى شكله البياني ولكن كثيراً ما نحتاج لتقدير درجة الالتواء بدقة فلا بد إذن من الاهتمام الى مقياس دقيق لهذه الظاهرة الهامة .

وسنقسم مقاييس الالتواء الى نوعين أساسيين ، الأول يصلح في جميع التوزيعات التي يمثل الوسط الحسابي نقطة التركيز فيها ، والثاني هو ما كان الوسيط أصح وأنسب في تمثيلها .

مقاييس الالتواء

أولاً - حالات يمثل الوسط الحسابي فيها التوزيع

إذا كان التوزيع التكراري متماثلاً فإن الوسط الحسابي والوسيط والمنوال تتساوى جميعها عند القيمة المقابلة لقمة المنحنى ، وكلما بعد التوزيع عن التماثل كلما اختلفت هذه المتوسطات الثلاث عن بعضها . وعلى ذلك يمكن أن نأخذ الفرق بين هذه المتوسطات كمقياس للالتواء . ولكن يكون من الأفضل أن نحصل على مقياس يمكن استخدامه في المقارنة ولهذا يقترح بيرسون أن نقسم الفرق بين الوسط والمنوال على الانحراف المعياري للتوزيع حيث أننا نعتبر الانحراف المعياري مقياس مدى ابتعاد القيم عن وسطها الحسابي .

$$\text{وبذلك يكون مقياس بيرسون للالتواء} = \frac{\text{الوسط الحسابي} - \text{المنوال}}{\text{الانحراف المعياري}}$$

فإذا كان التوزيع متماثلاً فإن الوسط الحسابي يتساوى مع المنوال وبذلك يكون مقياس الالتواء = صفر وكذلك معامل الالتواء . أما إذا كانت النتيجة أكثر من صفر دل ذلك على وجود التواء في التوزيع ودرجة هذا الالتواء تتوقف على قيمة المعامل وإشارته الجبرية . فإذا كان معامل الالتواء صغيراً دل ذلك على ضعف الالتواء والعكس إذا كبر المعامل فإن ذلك يدل على شدة الالتواء . وإذا كان الوسط الحسابي أكبر من المنوال فإن المعامل يكون موجباً وهذا دليل على التواء التوزيع نحو اليمين بمعنى أن القيم التي تزيد عن المنوال تكون تكراراتها أكبر من تكرارات القيم التي تقل عن المنوال . أما إذا كان الوسط الحسابي أقل من المنوال فإن المعامل يكون

سألباً وهذا دليل على التواء التوزيع نحو اليسار بمعنى ان القيم التي تزيد عن المتوسط تكون تكراراتها أقل من تكرارات القيم التي تقل عن المتوسط .

وعيب معامل الالتواء السابق أنه يعتمد على المتوسط وهو مقياس غير دقيق لا يجب أن نعتمد عليه في قياس درجة الالتواء . وخير الوسائل للاستغناء عنه التعبير عن الفرق بين الوسط الحسابي والمتوسط بدلالة الوسيط بدلاً من المتوسط . ويذكر القارئ ان هنالك علاقة تقريبية بين المتوسطات الثلاث في حالة التوزيعات المعتدلة الالتواء وهي :

$$\text{الوسط الحسابي} - \text{المتوسط} = 3 \text{ (الوسط الحسابي - الوسيط)}$$

ولكن المهم أن لا نستخدم هذه العلاقة إلا في حالة التوزيعات القريبة من التماثل أي المعتدلة الالتواء . ويكون معامل الالتواء في هذه الحالة :

$$\text{معامل الالتواء} = \frac{3 \text{ (الوسط الحسابي - الوسيط)}}{\text{الانحراف المعياري}}$$

على أنه يجب أن نلاحظ أن معامل الالتواء المحسوب على الأساس الأول يختلف بعض الشيء عن المعامل المحسوب على الأساس الثاني وذلك لأن العلاقة بين المقاييس علاقة تقريبية

ثانياً - حالات يمثل الوسيط فيها التوزيع

إذا كان التوزيع مفتوحاً فان بولي Bowley يقترح مقياساً آخر للالتواء وهو في نفس الوقت اسهل في الحساب من مقياس بيرسون Pearson . ويقوم مقياس بولي على أساس العلاقة بين الربع الأعلى والأدنى والوسيط، ذلك لأنه في حالة التوزيع التكراري المتماثل تتساوى المسافة بين كل من الربعين

والوسيط ، ولا يتحقق ذلك إذا كان التوزيع ملتوياً . وعلى ذلك يمكننا قياس الالتواء بطرح الفرق بين الربيع الأعلى والوسيط من الفرق بين الوسيط والربيع الأدنى . ذلك لأنه إذا كان التوزيع متاثلاً فإن الفرق يكون صفراً أما إذا كان الفرق أكبر من صفر فإن هذا دليل على وجود التواء في التوزيع تتوقف درجته ونوعه على قيمة الفرق وإشارته الجبرية ، فإذا كان الفرق كبيراً كان ذلك دليلاً على وجود التواء شديد والعكس إذا كان الفرق بسيطاً ، وإذا كان الفرق موجباً كان دليلاً على التواء التوزيع نحو اليمين والعكس إذا كان الفرق سالباً كان ذلك دليلاً على الالتواء نحو اليسار .

ولما كان هذا الفرق رقماً ميزافانه لا يصلح للمقارنة ولهذا يقترح بولي قسمة هذا الفرق على مجموع المسافة بين كل من الربيعين والوسيط أي على الانحراف الربيعي :

وعلى ذلك يكون مقياس بولي للالتواء = (الربيع الأعلى - الوسيط) - (الوسيط - الربيع الأدنى)

ويكون معامل بولي للالتواء =
$$\frac{\text{الربيع الأعلى} - \text{الوسيط} - \text{الربيع الأدنى}}{\text{الربيع الأعلى} - \text{الربيع الأدنى}}$$

وميزة هذا المقياس هو سهولة حساب الوسيط والربيعين بالنسبة لصعوبات حساب الانحراف المعياري والوسط الحسابي والمنوال .

على أنه يجب أن نلاحظ أن النتيجة التي نحصل عليها بتطبيق هذا المعامل تختلف عن تلك التي نحصل عليها إذا استخدمنا معامل بيرسون ، وذلك لأن كل من المقياسين يقوم على أساس مختلف ، فلا يجب استخدام المقياسين لمقارنة التواء توزيعين تكرارين وإنما يجب استخدام نفس المقياس إذا أردنا المقارنة .

العزوم ومقاييس الالتواء :

عزم أي قوة هو مقدار العمل الذي تحدثه ، ويتوقف هذا العمل على القوة نفسها والمسافة بين هذه القوة والنقطة التي عندها تحدث اثرها ،
فقوة مقدارها ٨ كيلو على بعد ١ قدم من نقطة الأصل (الصفر) تعادل في مفعولها قوة مقدارها ٢ كيلو على بعد ٤ قدم من نقطة الاصل ، أي أن التوازن يتحقق عند تساوي الناتج الموجب مع الناتج السالب .

وبالنسبة للتوزيع التكراري تكون تكرارات التوزيع هي القوى المؤثرة عليه وعزم أي تكرار يقاس بحاصل ضرب التكرار في انحرافه عن نقطة الأصل في التوزيع التي يعبر عنها الوسط الحسابي . والعزم الأول هو حاصل جمع عزوم تكرارات التوزيع مقسوماً على مجموع التكرارات في التوزيع أي $\frac{\sum x_k}{\sum k}$ ويعبر عنه في العزوم بالرمز M_1 وحيث أن

مجموع الانحرافات عن الوسط الحسابي = صفراً ، لذلك يكون العزم الأول في التوزيع التكراري مساوياً صفراً دائماً . أما إذا استبدلنا العزم الاول عن الوسط الحسابي بالعزم الاول عن القيمة صفر فمن الواضح أن الناتج يكون هو الوسط الحسابي نفسه .

$$\frac{\sum (x - \text{صفر})}{\sum k} = \text{أي ان العزم الأول (حول الصفر)}$$

= الوسط الحسابي .

لاحظنا أننا في هذه الحالة كأننا حسبنا الوسط الحسابي بدلالة وسط وسط فرضي صفر وبذلك يكون الوسط الحسابي مساوياً لمجموع الانحرافات عن الصفر مقسوماً على مجموع التكرارات (ارجع الى قانون الوسط الحسابي بدلالة الوسط الفرضي) .

والعزم الثاني للتوزيع التكراري وهو متوسط مربعات الانحرافات القيم في التوزيع (المراكز) عن الوسط الحسابي أي :

$$\sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

ومن الواضح أن العزم الثاني يعطينا التباين أي مربع الانحراف المعياري .

والعزم الثالث للتوزيع التكراري وهو متوسط مكعبات انحرافات القيم في التوزيع (المراكز) عن الوسط الحسابي أي :

$$\sigma^3 = \frac{\sum x^3}{n} - 3 \left(\frac{\sum x}{n} \right) \left(\frac{\sum x^2}{n} \right) + 2 \left(\frac{\sum x}{n} \right)^3$$

ويكون هذا العزم مساوياً صفر في التوزيع الطبيعي فقط ، أما في التوزيع اللتوي يكون العزم قيمة موجبة أو سالبة تبعاً لنوع الالتواء ، ذلك لأن تكعيب الانحرافات لا يخلصنا من الإشارة الجبرية فتظل انحرافات القيم الأقل من المتوسط تحمل إشارة مضادة لانحرافات القيم الأكبر عن المتوسط وبذلك عند الجمع يكون الناتج إما موجباً أو سالباً ، فإذا كانت الانحرافات الكبيرة سالبة يكون المجموع الجبري لمجموع مكعبات الانحرافات سالباً ولو كانت الانحرافات الكبيرة موجبة فإن المجموع الجبري لمكعبات الانحرافات يكون موجباً . وبذلك يمكن استخدام خاصية العزم الثالث في قياس درجة التواء التوزيع . وحيث أننا عند قياس الالتواء يكون هدفنا غالباً هو مقارنة التواء توزيعات مختلفة فلا يمكن الاعتماد على العزم الثالث في المقارنة ، ولذلك لكي نحصل على مقياس نسبي للالتواء نقسم العزم الثالث على مكعب الانحراف المعياري حيث أن وحدات القياس في البسيط تكون مكعبة ولذلك يجب القسمة على قيمة مرفوعة إلى نفس القوة في المقام .

$$\frac{3-m}{3} \text{ أي } \frac{3-m}{3(\sqrt{2})} = B_1 \text{ للتواء}$$

وقد اقترح بيرسون لأسباب رياضية استخدام مربع المعامل السابقة كمقياس للتواء :-

$$\frac{2(3-m)}{3(2)} = B_1 \text{ ولذلك يكون}$$

وحيث ان الانحرافات عن الوسط الحسابي يمكن ان تكون أرقام كبيرة وتكميها يؤدي الى ضياع وقت كبير ، لذلك بحسب العزم الثالث على أساس وسط فرضي ثم يصحح حتى نحوله الى العزم الثالث على أساس الوسط الحسابي باستخدام العلامة الآتية :-

$$3-m = \frac{\text{مح } 3 \text{ ك}}{\text{مح ك}} - 3 \times \frac{\text{مح } 2 \text{ ك}}{\text{مح ك}} + 2 \left(\frac{\text{مح } 1 \text{ ك}}{\text{مح ك}} \right)$$

مثال ٢٤ :

ف	ك	ح	ح ك	ح ٢ ك	ح ٣ ك
١٢	١	٣-	٣-	٩	٢٧-
١٤	٤	٢-	٨-	١٦	٣٢-
١٦	٦	١-	٦-	٦	٦-
١٨	١٠	صفر	صفر	صفر	صفر
٢٠	٧	١٠	٧+	٧	٧+
٢٢	٢	٢+	٢+		
٣٠			٨-	٤٢	٥٠-

$$3 - \frac{\text{م ح ك}^3}{\text{م ك}} - \frac{\text{م ح ك}^2}{\text{م ك}} \times \frac{\text{م ح ك}}{\text{م ك}} + \left(\frac{\text{م ح ك}}{\text{م ك}} \right)^2 = 3$$

$$= \frac{50}{30} - \frac{8}{30} \times 3 + \frac{42}{30} \times 2 + \left(\frac{80}{30} \right)^2 =$$

$$= 1.6667 - 0.8 + 2.8 + 7.1111 =$$

$$= 10.7778$$

وإذا أردنا معرفة العزم الثالث بوحدات طول الفئة تضرب النتيجة السابقة في مكعب طول الفئة على انه عند حساب الالتواء لا داعي لذلك حيث يمكن ان نقسم العزم الثالث على مكعب الانحراف المعياري (بدون الضرب في طول الفئة) ثم تربيع الناتج نحصل على معامل بيرسون للالتواء .

وبذلك يكون ع بدون التحويل الى وحدات القياس الخاصة بطول الفئة .

$$= \sqrt{\frac{42}{30} - \left(\frac{8}{30} \right)^2}$$

$$\therefore \text{ع}^3 = 1.0295$$

$$\therefore \text{معامل الالتواء} = \left(\frac{10.7778 - 1.0295}{1.0295} \right) = 9.47$$

وبالمثل يمكن حساب العزم الرابع حول الوسط الحسابي وهو عبارة عن متوسط مجموع انحرافات القيم عن الوسط الحسابي مرفوعة الى القوة الرابعة ويمكن التعبير عنه رمزيا

$$\frac{\text{م ح ء ك}}{\text{م ح ك}} = \text{م ء} = \text{م ح ك}$$

وعلى نفس النمط يمكن قياس العزم الخامس والسادس . . الخ ،
إلا أن العزوم الأربعة الأولى هي الشائعة الاستعمال في الاحصاء ، حيث أن
العزم الثاني يدلنا على التباين والعزم الثالث يدلنا على الالتواء والعزم الرابع
يدلنا على التفرطح . والتفرطح يعني ان قمة المنحنى الذي يمثل التوزيع
مفرطحة وليست مدببة ، ولا شك انه كلما كان تشتت الوحدات حول
متوسطها صغيراً كلما قل التفرطح والعكس كلما كان التشتت كبيراً كلما زاد
التفرطح ، إلا انه لا توجد أي صلة بين التشتت والتفرطح فكل منهما يصف
خاصية مختلفة تماماً في التوزيع التكراري فالتفرطح يُعنى به تفرطح قمة
المنحنى وليس انقراج المنحنى من أسفل أي اتساع المدى بين طرفيه على
المحور الافقي وهو ما يدل على تشتت التوزيع . والتوزيع الطبيعي بالرغم
من انه يكون عادة متوسط التفرطح الا أن درجة تفرطحه تختلف من منحنى
إلى آخر .

تمارين :

١ - احسب معامل الالتواء للتوزيع الآتي بالطرق المختلفة :

التكرار	الفئات
٢	٩٣ - ٩٧
٥	٩٨ - ١٠٢
١٢	١٠٣ - ١٠٧
١٧	١٠٨ - ١١٢
١٤	١١٣ - ١١٧
٦	١١٨ - ١٢٢
٣	١٢٣ - ١٢٧
١	١٢٨ - ١٣٢

٢ - إذا كان العزم الثاني لتوزيعين تكراريين هو ٩ ، ١٦ على التوالي ،
والعزم الثالث لنفس التوزيعين هو ٨ ، ١ - ٨ ، ١٢ ، أي من التوزيعين
أكثر التواء وما هو اتجاه التواء التوزيعين . التوزيع الاول أكثر التواء .



الفصل الثامن

توفيق المنحنى الطبيعي (المعتدل)

أشرت سابقاً إلى بعض خواص التوزيع الطبيعي (المعتدل) ، ونظراً لأهمية المنحنى في الدراسة الخاصة بالعينات يكون من المفيد ان ندرس خواصه دراسة تفصيلية .

لقد تبين لنا من الفصل السادس كيف يمكن تقسيم مساحة المنحنى الطبيعي الى قطاعات محددة بدلالة الوسط الحسابي والانحراف المعياري للتوزيع الذي يمثله هذا المنحنى . وبمعنى آخر يمكننا أن نعين نسبة مساحة المنحنى الطبيعي المحصورة بين قيمتين من قيم الظاهرة التي يمثّلها المنحنى إلى المساحة الكلية اذا عرفنا انحراف كل من القيمتين عن الوسط الحسابي مقدراً بوحدات الانحراف المعياري . فاذا كان انحراف كل من القيمتين عن الوسط الحسابي للتوزيع يساوي $+ 1$ ع أو $- 1$ ع ، فان المساحة المحصورة بين هاتين القيمتين تمثل 68.27% من المساحة الكلية للمنحنى / أي من مجموع الوحدات التي يمثّلها التوزيع . أما إذا كان انحراف كل من القيمتين عن الوسط الحسابي للتوزيع يساوي $+ 2$ ع أو $- 2$ ع ؛ فإن المساحة المحصورة بين هاتين القيمتين تمثل 95.45% من مجموع المساحة ، وإذا كان انحراف كل من القيمتين عن الوسط الحسابي للتوزيع يساوي $+ 3$ ع أو $- 3$ ع فان المساحة المحصورة بين القيمتين تمثل 99.73% من مجموع المساحة الكلية للمنحنى

على أن تقسيم مساحة المنحنى الطبيعي بدلالة الوسط الحسابي والانحراف المعياري لا يقتصر فقط على هذه القطاعات الثلاث الأساسية ، إذ يمكن التوسع في تطبيق هذه الخاصة بأعداد جدول مساحات المنحنى الطبيعي ومنه يمكن تحديد نسبة المساحة المحصورة بين أي قيمة وبين خط التآكل من مجموع المساحة الكلية إذا استطعنا تحديد انحراف هذه القيمة عن الوسط الحسابي للتوزيع بدلالة وحدات الانحراف المعياري ، ويكون ذلك بقسمة الفرق بين القيمة وبين الوسط الحسابي على الانحراف المعياري وبذلك نستطيع أن نعرف مقدار هذا الفرق بدلالة وحدات الانحراف المعياري ، أي هل يساوي الفرق ± 1 ع أو ± 1.2 ع أو ± 1.5 ع أو ± 1.7 ع أو ± 2.2 ع ، فإذا حددنا ذلك نستطيع أن نحدد المساحة المطلوبة كنسبة من المساحة الكلية .

وحيث أن كثيراً من الدراسات الإحصائية تقوم على أساس المعاينة وحيث أن نتائج العينة لا يمكن أن تطابق تماماً صفات المجتمع الإحصائي الذي سحبت منه العينة ، لذلك قد يحدث أحياناً أنه بعد تبويب البيانات الخاصة بالعينة في توزيع تكراري يترأى لنا أن هناك ما يدعو إلى الاعتقاد بأن هذه البيانات موزعة توزيعاً طبيعياً خاصة إذا أعدنا التجربة على عينات أخرى ووجدنا أنه في كل مرة توزع قيم الظاهرة توزيعاً يقرب من الاعتدال وأنه إذا كبرنا حجم العينة قرب التوزيع من الوضع المعتدل . وبذلك يمكن في هذه الحالة أن نستنتج أنه لو كان من الممكن عملياً جمع البيانات من جميع وحدات المجتمع الإحصائي موضوع الدرس لوجدنا أنها موزعة فعلاً توزيعاً معتدلاً . ولما كانت دراسة العينات هي الوسيلة لدراسة المجتمعات التي سحبت منها فيمكن أن نكتفي بعينة واحدة على أن نمهد المنحنى الطبيعي الذي يمثل التوزيع أي نجعله يتخذ الوضع المثالي الذي يكون عنده معتدلاً تماماً . ولزيادة الإيضاح نقول أن دراسة عينة من مجتمع ما لا يمكن أن يظهر فيها

نفس التوزيع النسبي لمختلف أنواع الوحدات التي تكوّن هذا المجتمع ، فبعض الأنواع قد تظهر في العينة بوزن يزيد عن وزنها في المجتمع ، ولهذا إذا كانت توزيع وحدات المجتمع توزيعاً طبيعياً فإن ذلك لن يظهر تماماً في العينة وبمعنى آخر إذا رسمنا المنحنى التكراري الذي يمثل وحدات العينة فلن يكون شكله معتدلاً مائة في المائة وإنما ستظهر فيه بعض التعرجات الناتجة عن ظهور بعض الوحدات في العينة بنسبة أكبر من نسبتها في المجتمع وظهور البعض الآخر بنسبة أقل من نسبتها في المجتمع . إلا أن هذه التعرجات في المنحنى قد تكون من الصغر بحيث يمكن أن نحكم بأن المنحنى لو ظهرت وحدات العينة بنفس توزيعها النسبي في المجتمع يمكن أن يكون منحنى تام الاعتدال ولهذا نقول « كيف يمكن أن يكون شكل المنحنى الذي يمثل العينة لو ظهرت الوحدات فيها بنفس توزيعها النسبي في المجتمع الذي تمثله ؟ وهذا القول هو الذي يدعونا الى توفيق المنحنى الذي يمثل وحدات العينة أي القضاء على التعرجات الشاذة التي تظهر فيه .

وعند توفيق المنحنى الذي يمثل وحدات العينة لا بد ان يتفق المنحنى بعد التوفيق مع المنحنى قبل التوفيق في المجموع الكلي للوحدات وفي الوسط الحسابي لقيمتها وكذلك في انحراف هذه القيم عن الوسط الحسابي أي في انحرافها المعياري . فنحن وإن كنا نغيّر من شكل المنحنى أي من توزيع الوحدات في الفئات المختلفة فلا يجب أن يتغيّر بذلك مجموع التكرارات أو الوسط الحسابي أو الانحراف المعياري للتوزيع الأصلي الذي نقوم بتوقيفه لأن تغيير هذه المقاييس يعني تغيير التوزيع تغييراً جذرياً وهذا ليس المقصود من عملية التوفيق ولهذا تؤخذ هذه المقاييس كدعائم يرتكز عليها عملنا الحسابي الذي تتضمنه عملية التوفيق .

والمقصود بعملية التوفيق هو حساب التكرارات النظرية التي تقابل التكرارات الواقعية أي المشاهدة نتيجة جمع المعلومات الاحصائية عن الظاهرة

موضوع الدرس . ويمكن إجراء عملية التوفيق بطريقتين ، اما عن طريق خاصية التوزيع الطبيعي التي سبق الاشارة اليها وهي امكانية تقسيم مساحته الى قطاعات معلوم نسبتها الى المساحة الكلية وفي هذه الحالة نستخدم جدول مساحات المنحنى الطبيعي ، واما عن طريق خاصية اخرى وهي إمكانية تحديد الاحداثيات الرأسية للمنحنى المعتدل .

توفيق المنحنى بطريقة المساحات :

اشرت في الفقرة السابقة الى الملاحظتين لآتينين : -

١ - ان عملية التوفيق لا يجب أن تغير من مجموع التكرارات أو الوسط الحسابي أو الانحراف المعياري للتوزيع الأصلي ولهذا نحسب هذه المقاييس أولاً وتؤخذ كدعائم يرتكز عليها عملنا الحسابي .

٢ - إن تطبيق هذه الطريقة يحتاج الى استخدام جدول مساحات المنحنى الطبيعي ، وهذا الجدول يعطي نسبة المساحة المحصورة بين الاحداثي الرأسي المتوسط للتوزيع (أي الذي يقابل قمة المنحنى والوسط الحسابي على المحور الأفقي) وبين الاحداثي الرأسي المقام على أي نقطة تمثل أي قيمة على يمين أو على يسار الوسط الحسابي للتوزيع . والجدول يعطي هذه النسبة إذا علمنا نسبة الفرق بين القيمة والوسط الحسابي للتوزيع الى الانحراف المعياري له أي

$$\frac{س - س-}{ع}$$
 وبمعنى آخر إذا علمنا كم يساوي الفرق بين القيمة والوسط

الحسابي من وحدات الانحراف المعياري للتوزيع . ويجب أن نلاحظ أن الجدول يعطي هذه النسبة على أساس ان المجموع الكلي لمساحة المنحنى واحد صحيح ؛ فاذا أردنا هذه النسبة في كل شكل مثوي فما علينا إلا أن نضرب النسبة التي تظهر في الجدول في ١٠٠ . كذلك يجب أن نلاحظ أن الجدول يعطي النسبة على أساس وجود القيمة على يمين أو على يسار محور التماثل .

مثال ٢٥ :

وفيما يلي اورد المثال التالي لايضاح طريقة العمل :

حسبت الأجور اليومية لألف عامل وبوبت في الجدول الآتي :

عدد العمال	فئات الأجر بالليرة
٦	٤ر٥ -
٧	٥ر٥ -
٣٥	٦ر٥ -
٤٨	٧ر٥ -
٦٥	٨ر٥ -
٩٠	٩ر٥ -
١٣١	١٠ر٥ -
١٧٣	١١ر٥ -
١٥٥	١٢ر٥ -
١١٧	١٣ر٥ -
٧٥	١٤ر٥ -
٥٢	١٥ر٥ -
٢١	١٦ر٥ -
٩	١٧ر٥ -
٦	١٨ر٥ - ١٩ر٥

واقترح احد الاحصائيين أن هذه البيانات (على أساس عينة من ١٠٠٠

عامل) توحى بأن التوزيع الذي يمثل هذه الظاهرة هو توزيع طبيعي ،
ولذلك نبدأ بتوفيق المنحنى الطبيعي أي بحساب تكرار كل فئة لو كان
التوزيع طبيعياً تماماً .

خطوات العمل :

١ - حساب الوسط الحسابي والانحراف المعياري للتوزيع الواقعي .

٢ - حساب الفرق بين كل قيمة (س) تمثل الحدود الدنيا للفئات والوسط
الحسابي للتوزيع .

٣ - حساب نسبة هذه الفروق الى الانحراف المعياري أي حساب

$$\text{الدرجة المعيارية} = \frac{س - س^-}{ع} .$$

٤ - استخراج نسبة المساحة المحصورة بين كل قيمة ومحور التماثل في
التوزيع من جدول المساحات للتوزيع الطبيعي .

٥ - طرح كل نسبة من التي تليها لتحديد نسبة مساحة كل فئة الى
المجموع الكلي للمساحة .

٦ - لما كانت المساحة الكلية للتوزيع يمثلها مجموع التكرارات ، لذلك
تضرب نسبة مساحة كل فئة في مجموع التكرارات نحصل على
التكرار النظري لكل فئة .

ويمكن اجراء الخطوات السابقة في الجدول الآتي -

لغات	ك	م	ح	ح ك	ح آك	محت	م-م ع	المساحة بين س و د المتماثل ب	مساحة الفئة ب	تكرار الفئة
مر-٤	٦	٥	٧-	٤٢	٢٩٤	٧٥٠٦	٢٨٦	٤٩٧٩	٤٥	٦
مر-٥	١٧	٦	٦-	١٠٢-	٦١٢	٦٥٠٦	٢٤٨	٤٩٣٤	١٣	١٢
مر-٦	٣٥	٧	٥-	١٧٥-	٨٧٥	٥٥٠٦	٢١	٤٨٢١	٢٥٧	٢٦
مر-٧	٤٨	٨	٤-	١٩٢-	٧٦٨	٤٥٠٦	١٧١	٤٥٦٤	٤٨١	٤٨
مر-٨	٦٥	٩	٣-	١٦٥-	٥٨٥	٣٥٠٦	١٣٣	٤٥٨٣	٢٦٤	٧٩
مر-٩	٩٠	٤	٢-	١٨٠-	٣٦٠	٢٥٠٦	٩٥	٢٢٨٩	١١٣٢	١١٣
مر-١٠	١٣١	١١	١-	١٣١-	١٣١	١٥٠٦	٥٧	٢١٥٧	١٤٠٤	١٤٠
مر-١١	١٧٣	١٢	صفر	صفر	صفر	٥٠٦	١٩	٧٥٣-	١٥٠٦	١٥١
مر-١٢	٤٥٥	١٣	١	١٥٥	١٥٥	٤٩٤	١٩	٧٥٣ +	١٤٠٤	١٤٠
مر-١٣	١١٧	١٤	٢	٢٣٤	٤٦٨	١٤٩٤	٥٧	٢١٥٧	١١٣٢	١١٣
مر-١٤	٧٥	١٥	٣	٢٢٥	٦٧٥	١٤٩٤	٩٥	٢٢٨٩	٧٩٤	٧٩
مر-١٥	٥٢	١٦	٤	٢٠٨	٨٣٢	١٤٩٤	١٣٣	٤٥٨٣	٤٨١	٤٨
مر-١٦	٢١	١٧	٥	١٠٥	٥٢٥	١٤٩٤	١٧١	٤٥٦٤	٢٥٧	٢٦
مر-١٧	٩	١٨	٦	٥٤	٣٢٤	٥٠٦	١٠٩	٤٨٢١	١١٥	١٢
مر-١٨	١٩	١٩	٧	٤٢	٢٩٤	١٤٩٤	٢٤٧	٤٩٣٢	٤٦	٦
مر-١٩	١				٧٤٩٤	٢٨٥	٢٨٥	٤٩٧٨		

$$٦٨٩٨ \quad ٦ + \quad ١٠٠٠$$

$$س- = ١٢ + ١ \frac{٦}{١٠٠٠} = ١٢.٠٠٦ ليرة$$

$$ع = ١ \sqrt{\frac{٢٦}{١٠٠٠} - \frac{٦٨٩٨}{١٠٠٠}}$$

$$= ١ \sqrt{٢٨٩٨ - ٠.٣٦٠٠٠٠}$$

$$= ١ \sqrt{٢٨٩٧٩٦٤}$$

$$= ٢٦٢٦ ليرة .$$

حصلنا في العمود الأخير من الجدول السابق على التكرارات النظرية أي للتكرارات التي يمكن توقعها لو كان التوزيع ممثلاً تمام التماثل . وبالقاء نظرة على هذه التكرارات يتبين أن هناك فروق بينها وبين التكرارات الواقعية وهو أمر كنا نتوقعه . والمشكلة هي - هل هذه الفروق من الكبر بحيث يتعين علينا استبعاد الافتراض الذي بدأنا به وهو اعتدال المجتمع الذي يمثله هذا التوزيع أم انها من الصغر بحيث يمكن أن نعزوها إلى المصادفة الناتجة عن المعاينة العشوائية . والبت في هذه المشكلة يكون بإجراء اختبار كاي تربيع الذي سنأتي إلى مناقشته في الفصل التالي :

توفيق المنحنى بطريقة الاحداثيات الرأسية :

استطاع الرياضيون أن يحددوا المعادلة التي تمثل المنحنى المعتدل وهي كالآتي :

$$ص = \frac{1}{\sqrt{ع} \cdot ط} \cdot \frac{س - ع}{ع^2}$$

حيث ص هي الاحداثي الرأسي المناظر لقيمة س .

، س هي الاحداثي الافقي مقيساً من نقطة الأصل التي تقع عند الوسط الحسابي للتوزيع وبعبارة أخرى فان قيم س تمثل الانحرافات عن الوسط الحسابي للتوزيع :

، ع هي الانحراف المعياري للتوزيع

، ط هي النسبة التقريبية (٣١٤١٦)

، ه هي أساس لوغاريتمات نابيير (٢٧١٨٢٨) .

وإذا أردنا تحديد الاحداثي الرأسي لتوزيع طبيعي معين عند أي قيمة على المحور الافقي يجب ضرب المعادلة السابقة في عدد وحدات التوزيع ، وبمعنى آخر يستبدل العدد بوحدات التوزيع . وبذلك يكون الاحداثي الرأسي المتوسط للتوزيع أي الاحداثي المقام على الوسط الحسابي للتوزيع =

$$\frac{1}{\text{ع} \sqrt{2} \text{ط}} \quad \text{وإذا كان مجموع التكرارات} = \text{م ك} \quad \text{يكون الاحداثي}$$

$$\frac{\text{م ك}}{\text{ع} \sqrt{2} \text{ط}} = \text{الرأسي المتوسط} \quad \text{حيث أن} \quad \frac{\text{س}^2 - \text{ع}^2}{\text{ع}^2} = \frac{\text{س}^2 - \text{ع}^2}{\text{ع}^2}$$

سوف تساوي صفر لأن القيمة - س = صفر إذ أن انحرافها عن الوسط الحسابي ، يساوي صفر .

مثال ٢٦ :

منحنى معتدل يمثل ١٠٠٠ وحدة ، وانحرافه المعياري = ١٠ كم يكون التكرار الذي يدل عليه محور التماثل فيه :

$$\text{ص} = \frac{1000}{31416 \times 2 \sqrt{10}} = 39894$$

وبذلك يمكن تحديد أي احداثي رأسي للمنحنى بضرب محور التماثل فيه في القيمة :

فلاحداثي الرأسي فوق النقطة التي تبعد عن الوسط

$$\frac{400}{200} -$$

الحسابي بمقدار ٢٠ = ٣٩٨٨٩٤ × ٢٧١٨٢٨

$$. ٥٣٩٩ = \frac{1}{2 \times ٢٧١٨٢٨} \times ٣٩٨٨٩٤ =$$

ولتسهيل العمل عند تحديد الاحداثيات الرأسية التي تكون المنحنى المعتدل يمكن الرجوع إلى جدول احداثيات المنحنى الطبيعي . ويعطى هذا الجدول ارتفاع الاحداثي الراسي عند أي نقطة تبعد عن الوسط الحسابي للتوزيع كنسبة مئوية من الاحداثي الراسي المتوسط (خط التماثل) . والابتعاد عن الوسط الحسابي يكون بدلالة الانحراف

$$. \frac{س - س^-}{ع} \text{ المعيارى للتوزيع أي}$$

لذلك حق يمكن استعمال هذا الجدول يجب تحديد الابتعاد عن الوسط الحسابي للتوزيع بدلالة الانحراف المعيارى أي تحديد قيمة $\frac{س - س^-}{ع}$ ،

والقيمة المقابلة في الجدول (أي القيمة المقابلة للدرجة المعيارية $\frac{س - س^-}{ع}$) تضرب

في الاحداثي الواسي المتوسط :

$$\frac{س ك}{ع \times ٢٢٤١٦} \text{ ويحدد هذا الاحداثي باستخدام المعادلة}$$

$$\text{أي : } \frac{س ك}{ع \times ٢٥٠٦٦٢٨}$$

ويمكن وضع هذه المعادلة في الصيغة :

$$\frac{\text{م ك} \times \text{طول الفئة}}{\text{ع (بالوحدات الأصلية للتوزيع)}} \times ٠.٣٩٨٩٤$$

مثال ٢٧ :

منحنى طبيعي يمثل ٦٠٠ وحدة وانحرافه المعياري = ٠.٠٠٨٥ ووسطه الحسابي = ٠.٠٢٠٢ ، كم يكون المحور الرأسي للقيمة ٠.٠١٨٢ إذا كان طول الفئة في هذا التوزيع = ٠.٠٠٠٤ ؟

$$\text{س} - \text{س} = ٠.٠١٨٢ - ٠.٠٢٠٢ = ٠.٠٠٢$$

$$\frac{\text{س} - \text{س}}{\text{ع}} = \frac{٠.٠٠٢}{٠.٠٠٠٨٥} = ٢.٣٥$$

من جدول الاحداثيات الرأسية للمنحنى الطبيعي يكون الاحداثي الرأسي المقابل لهذه القيمة بشكل ٦٢١٪ من الاحداثي المتوسط .
والخطوة التالية هي حساب الاحداثي الرأسي المتوسط باستخدام المعادلة :

$$١١٢.٦ = \frac{٦٠٠ \times ٠.٠٠٠٤}{٠.٠٠٠٨٥} \times ٠.٣٩٨٩٤$$

أو باستخدام المعادلة :

$$١١٢.٦ = \frac{٦٠٠}{\sqrt{٢.١٢٥ \times ٣.١٤١٦}}$$

ونلاحظ ان القيمة ٢.١٢٥ هي الانحراف المعياري بدلالة طول الفئة في التوزيع أي بقسمة الانحراف المعياري بالوحدات الأصلية للتوزيع على طول الفئة .

وبذلك يكون الاحداثي الرأسي للقيمة ٠.٠٠٨٤ =

$$6321 \times \frac{1126}{100} = 712$$

مثال ٢٨ :

ويمكن توضيح عليه توفيق المنحنى بطريقة الاحداثيات الرأسية بالجدول الآتي :

الفئات	ك	م	م - س	م - س - ع	نسبة الاحداثي الرأسي إلى الاحداثي المتوسط	احداثي المركز
٤٥ -	٦	٥	٧٠٠٦	٢٦٦	٠٢٩٠٨	٦
٥٥ -	١٧	٦	٦٠٠٦	٢٣٣	٠٦٦٢٤	١٢
٦٥ -	٣٥	٧	٥٠٠٦	١٩١	١٦١٣٧	٢٥
٧٥ -	٤٨	٨	٤٠٠٦	١٥٢	٣١٥٠٠	٤٨
٨٥ -	٦٥	٩	٣٠٠٦	١١٤	٦٢٢١٤	٧٩
٩٥ -	٩٠	١٠	٢٠٠٦	٧٦	٧٤٩١٦	١١٣
١٠٥ -	١٣١	١١	١٠٠٦	٣٨	٩٣٠٢٤	١٤١
١١٥ -	١٧٣	١٢	٠٠٠٦	صفر	١٠٠٠٠	١٥٢
١٢٥ -	١٥٥	١٣	٠٩٩٤	٣٨	٩٣٠٢٤	١٤١
١٣٥ -	١٦٧	١٤	١٩٩٤	٧٦	٧٤٩١٦	١١٣
١٤٥ -	٧٥	١٥	٢٩٩٤	١١٤	٥٢٢١٤	٧٩
١٥٥ -	٥٢	١٦	٣٩٩٤	١٥٢	٣١٥٠٠	٤٨
١٦٥ -	٢١	١٧	٤٩٩٤	١٩١	١٦١٣٧	٢٥
١٧٥ -	٩	١٨	٥٩٩٤	٢٣٣	٠٦٦٢٣	١٢
١٨٥ -	٦	١٩	٦٩٩٤	٢٦٦	٠٢٩٠٨	٦

$$\text{الاحدائي الرأسى المتوسط} = \frac{1 \times 1000}{27626} \times 39894 = 1519$$

وتكون الخطوة التالية إجراء اختبار كاي تربيع للحكم على حسن المطابقة بين التكرارات الواقعية والتكرارات النظرية التي استنتجناها من عملية التوفيق .

لماذا نوفق المنحنى المعتدل :

يمثل التوزيع التكراري عادة عينة مسحوبة من مجتمع معين على أساس ان تمثل صفات هذا المجتمع الذي انتخبت منه العينة ، وحيث انه لا يمكن دراسة جميع وحدات المجتمع الذي تمثله العينة فاننا نستنتج هذه الصفات من دراستنا للعينة ، ولهذا نقوم بعملية توفيق المنحنى المعتدل أو أي منحنى آخر محاولين أن نعرف كيف يكون شكل المنحنى الذي يمثل المجتمع كله نظراً لأن المنحنى الذي يمثل العينة قد يكون به بعض التمرجات الناشئة عن ظهور بعض الوحدات الشاذة بوزن يفوق وزنها في المجتمع . ونحن عندما نوفق المنحنى المعتدل لتوزيع خاص بعينة ما فاننا بذلك نبدي رأياً ، وطبعاً لا يكون رأياً اعتباطياً وإنما يكون قلناً على اساس ما نشاهده من بعض التائل في التوزيع خاصة اذا كان معامل الالتواء الذي نحسبه صغيراً جداً . ولا يجب ان نتوقف عند إبداء الرأي فقط بل نقوم بعملية التوفيق حتى نستنتج ما نتوقعه من تكرارات كي يمكن ان نجري عليها اختبار حسن المطابقة (اختبار كاي تربيع) الذي يمكننا من تأكيد الرأي او استبعاده .

كذلك نستطيع بتوفيق المنحنى المعتدل ان نحدد تعميماً على المجتمع الذي تمثله العينة نسبة الوحدات التي نتوقع ان تكون أقل أو أكثر أو بين قيمتين معينتين . مثلاً اذا وفقنا المنحنى الذي يمثل طول عمر اللبات الكهربائية ان نعرف تعميماً على مجموع اللبات الكهربائية المنتجة نسبة اللبات التي سوف تستمر في

الاضاءة ١٥٠٠ ساعة أو أكثر أو أقل من هذا الزمن أو غيره . كذلك يمكن بتوفيق المنحنى الخاص بتوزيع يمثل الوفيات حسب فئات العمران نعرف تعميما نسبة عدد الافراد الذين يتوفون في كل فئة عمر . كذلك نستطيع بتوفيق المنحنى المعتدل ان نحدد على اساس البيانات التي نأخذها من عينة ما خاصة بمقاييس الرجال او النساء او الاطفال نسبة الوحدات التي يمكن انتاجها من كل نوع من الملابس تبعاً للمقاييس المختلفة التي نريدها .

استعمال التوزيع المعتدل في اختبار الفروض .

على اساس توزيع ظاهرة ما توزيعاً معتدلاً نستطيع ان نحدد ما اذا كانت وحدة ما بالنسبة لهذه الظاهرة هي وحدة عادية او غير عادية (شاذة) وفي الغالب تكون الوحدة عادية اذا كانت ضمن الـ ٩٥ ٪ ، ومشكوكا في امرها اذا كانت واقعة بين ٩٥ ٪ و ٩٩ ٪ ، وشاذة اذا كانت خارج الـ ٩٩ ٪ ومن الواضح ان تحديد وقوع الوحدة في اي قطاع من هذه القطاعات يتوقف على بعد قيمتها عن الوسط الحسابي محسوبة بالدرجات المعيارية اي نسبة الفرق الى الانحراف المعياري للتوزيع .

مثال ٢٩ :

اذا كان متوسط اجر العامل في احدى الصناعات ٣٥ ليرة في الاسبوع ، بانحراف معياري ٩ ليرات . ووجد ان عامل ما يتقاضى ٦٠ ليرة في الاسبوع هل اجر هذا العامل يختلف عن اجور عمال هذه الصناعة .

$$\text{الدرجة المعيارية} = \frac{35 - 60}{9} ، \frac{35}{9} : 278 \text{ وهو اكبر من مستوى}$$

معنوية ٢٥٨ المقابل لدرجة ثقة ٩٩ ٪ . وبذلك نرفض الفرض بان هذا العامل لا يختلف اجره عن اجور العمال في الصناعة . أي ان هناك اختلاف

جوهري بين اجره والاجور عامة في هذه الصناعة ويمكن الاجابة على السؤال بطريقة اخرى .

$$\begin{aligned} \text{حدي الفئة بمستوى معنوية } 99\% &= 9 \times 2058 + 35 \\ \text{وبمستوى معنوية } 95\% &= 9 \times 1996 + 35 \end{aligned}$$

ونلاحظ ان الاجر ٦٠ يقع خارج هذا المدى ، وبذلك يكون هناك فرق جوهري بين اجر العامل والاجور عامة في هذه الصناعة .

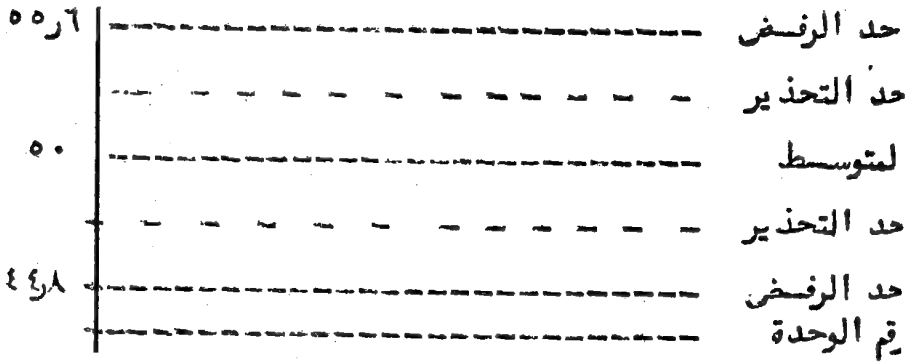
وعلى اساس هذه الفائدة للتوزيع المعتدل يمكن ضبط جودة الانتاج في مصنع ما حيث نستطيع أن نتعرف على الوحدات التي يكون انتاجها شاذاً بالنسبة للمتوسط العام الخاص بانتاج هذا المصنع .

مثال ٣٠ - يقوم أحد المصانع بانتاج الاسمنت وقد ضبطت آلات التعبئة بحيث تملأ ٥٠ كلغ في كل كيس بانحراف معياري ٢ كجم والمطلوب رسم حدود ضبط الانتاج .

$$\begin{aligned} \text{بمستوى معنوية } 95\% \text{ يكون حدي الثقة } &= 2 \times 1996 + 50 \\ &= 5392 \text{ كجم ، } 4618 \text{ كجم} \end{aligned}$$

$$\begin{aligned} \text{وبمستوى معنوية } 99\% \text{ يكون حدي الثقة } &= 2 \times 2058 + 50 \\ &= 5516 \text{ كجم ، } 4484 \text{ كجم} \end{aligned}$$

وبذلك يمكن رسم خريطة لمراقبة الانتاج على أساس هذه الحدود للثقة كالاتي : -



فاذا عبثت الآلة الأكياس بالاوزان الآتية ٥١ ، ٥٢٦ ، ٥٠ ، ٥٤٩ ، ٥١٨ ، ٥٧ ، ٩٥٠ كجم ، نلاحظ أن السابعة مرفوضة والثانية والخامسة مشكوك في امرهما وباقي الأكياس مقبولة .

ومن الواضح اننا في الاختبار السابق تعتبر الوحدة شاذة اذا كانت مفرطة في الزيادة أو مفرطة في النقص . فاذا كنا نريد أن نجيب على السؤال ما اذا كانت الوحدة مفرطة في الزيادة فقط أو مفرطة في النقصان فقط ، فاننا في هذه الحالة نهتم بناحية المنحنى التي تتعلق بالاختبار ، فاذا كنا نسأل هل الوحدة مفرطة في الزيادة لا يهمنا الجانب الأيسر من المنحنى ، واذا كنا نسأل هل الوحدة مفرطة في النقصان لا يهمنا الجانب الايمن من المنحنى . فاذا كنا نهتم بالجانب الايمن نضيف ٢٥٨ع الى الوسط الحسابي ، واذا كان اهتمامنا موجهاً الى الجانب الايسر في المنحنى نطرح ٢٥٨ع من الوسط الحسابي :

تـمـارـين

١ - وفق المنحنى المعتدل للتوزيع الآتي بطريقتين :

ك	ف
١	-٨٠
٨	-٨١
٣٥	-٨٢
٨٢	-٨٣
١٢٢	-٨٤
١٢٤	-٨٥
٨٣	-٨٦
٣٤	-٨٧
١٠	-٨٨
١	-٨٩ - ٩٠

(التكرارات بعد التوفيق هي ٢ ، ٩ ، ٣٤ ، ٨١ ، ١٢٤ ، ١٢٤ ، ٨١ ، ٣٤)
(٣٤ ، ٩ ، ٢)

٢ - في توزيع طبيعي ٧٪ من عدد القيم أقل من ٣٥ ، ٨٩٪ من هذه القيم أقل من ٦٣ . ما هو الوسط الحسابي والانحراف المعياري لهذا التوزيع .

$$س = -٥٠.٣ \quad ع = ١٠٣.٩$$

٣ - اذا كان متوسط درجات الطلبة ٧٢ بانحراف معياري ٩ - فاذا تقرر أن يعطى ١٠٪ من الطلبة درجة ممتاز ، ما هي أقل درجة يجب أن يحصل عليها هؤلاء الطلبة (الدرجة = ٨٤) .

٤ - اذا كان احد المصانع ينتج انابيب متوسط قطرها ٠.٦١٤٠ بوصة بانحراف معياري ٠.٢٥ بوصة . احسب نسبة الانابيب التي ينتجها المصنع بالمقاييس الآتية :

أ - بين	٠.٦١	بوصة	٠.٦١٨	=	٩٣	%
ب - اكبر من	٠.٦١٧	بوصة		=	٨١	%
ج - أقل من	٠.٦٠٨	بوصة		=	٤٧	%
و - قطرها يساوي	٠.٦١٥	بوصة		=	١٥	%

٥ - ينتج احد المصانع انابيب متوسط قطرها ٠.٥٠٢ بوصة بانحراف معياري ٠.٠٥ بوصة فاذا كانت الانابيب تقبل اذا كان قطرها يتراوح بين ٠.٤٩٦ بوصة و ٠.٥٠٨ بوصة احسب نسبة الانابيب المرفوضة .

(النسبة = ٢٣ %)

ارسم لوحة لمراقبة الانتاج في هذا المصنع وفسر كيف يمكن استعمالها .

٦ - اذا كان متوسط الاوزان لخمسة طالب هو ١٥١ رطل بانحراف معياري ١٥ رطل كم عدد الطلبة الذين يتراوح اوزانهم بين ١٢٠-١٥٥ رطل (٣٠٠ طالب) . وكم عدد الطلبة الذين تزيد اوزانهم عن ١٨٥ رطل (٥ طلاب) .

الفصل التاسع

اختبار كباى تربيع ⁽²⁾

يواجه الاحصائي مشكلة تتخذ صوراً مختلفة في كثير من الدراسات الاحصائية ، مشكلة اختبار التطابق بين بيانات واقعية وبيانات مستنبطة على اساس افتراض معين أو على أساس اجراء الدراسة بالمعيار حيث نحصل على بيانات من العينة تختلف بعض الشيء عن المتوقع طبقاً لقوانين الاحتمالات، حتى يمكن بذلك ان يحكم على افتراضه فاما أن يطمئن اليه ويعتبره مناسباً غير بعيد عن الحقيقة واما أن لا يعتبره كذلك فيستبعده . واختبار كباى تربيع هو الاداة الاحصائية التي يمكن بواسطتها مواجهة هذه المشكلة والبت فيها . وللمناقشة طريقة اجراء هذا الاختبار يكون من الأفضل أن أورد أمثلة ملموسة توضح خطوات العمل اللازمة .

مثال ٣٠ :

في عام ١٩٥٠ قام مركز الدراسات الاستقصائية باحدى الجامعات الاميركية بدراسة حول العلاقة بين الحالة الزوجية لرب الاسرة والادخار ؛ ولقد تبين من المعلومات التي جمعت عن ٣٣٢٧ أسرة الوضع التالي فيما يختص بموضوع الدراسة :

الحالة الزوجية لرب الاسرة	عدد المدخرين	عدد غير المدخرين	المجموع
اعزب	٤٩٠	٣٩٠	٨٨٠
متزوج	١٥٥٢	٧٩٥	٢٤٤٧
المجموع	٢٠٤٢	١٢٨٥	٣٣٢٧

من هذه المعلومات يرغب الاحصائي في الاجابة على السؤال الذي يتبادر الى الذهن في هذه الحالة وهو هل هناك اختلاف جوهري بين سلوك العزاب والمتزوجين نحو الادخار او ان الاختلاف بين سلوكهم ليس الا اختلاف ظاهري يمكن ان يكون راجعاً لمجرد الصدفة الناتجة من دراسة عينة من الاسر جاءت نتائجها معطية بعض الوزن (عشوائياً) لاحدى الفئات الاربع التي يتكون منها الجدول السابق

للاجابة على هذا السؤال يبدأ الاحصائي بافتراض ان لا فرق هناك بين سلوك كل من الفريقين ، أي انها يكونا مجتمعاً واحداً سلوكه متشابه نحو الادخار . وبمعنى آخر يفترض الاحصائي أن لا علاقة البتة بين الحالة الزوجية لرب الاسرة وسلوكه نحو الادخار اي لا توجد أية علاقة بين تصنيف أرباب الاسر الى مدخرين وغير مدخرين وتصنيفهم الى عزاب ومتزوجين . على هذا

الأساس تكون النسبة العامة للمدخرين في هذا المجتمع المفترض هي $\frac{3242}{3327}$

(مجموع المدخرين في العينة الى المجموع الكلي لوحداتها)

على اساس هذه النسبة العامة وتبعاً للافتراض بأن كل من الفريقين ليس الا صورة ينعكس عليها السلوك العام في المجتمع انعكاساً كاملاً يكون عدد المدخرين وغير المدخرين في كل من الفريقين كالآتي :

$$\left. \begin{aligned} & 880 \times \frac{3242}{3327} = 840 \text{ مدخرين} \\ & 880 - 840 = 40 \text{ غير مدخرين} \end{aligned} \right\} \text{العزاب}$$

$$\left. \begin{aligned} & 2447 \times \frac{3242}{3327} = 2412 \text{ مدخرين} \\ & 2447 - 2412 = 35 \text{ غير مدخرين} \end{aligned} \right\} \text{المتزوجون}$$

لما كانت الأرقام المحسوبة للمدخرين وغير المدخرين في كل من الفريقين قد أجري حسابها على أساس الافتراض الذي بدأنا به ، لذا نسميها بالتكرارات النظرية أي الأرقام التي كان يمكن ان تظهر في الجدول السابق لو كانت الحالة الزوجية لرب الأسرة وسلوكه نحو الادخار ظاهرتان مستقلتان عن بعضها اي لا علاقة البتة بينهما . وباستعراض هذه التكرارات النظرية وبمقارنتها مع التكرارات الواقعية الواردة في الجدول الأصلي نلاحظ وجود فرق بينها ، والمشكلة التي تواجهنا بعد ذلك هي كيف يمكن ان نحكم على هذا الفرق لنقرر ما إذا كان فرقاً بسيطاً يمكن أن يكون راجعاً إلى الصدفة أو انه من الكبر بحيث لا يمكن ان يكون سببه مجرد الصدفة . ولا شك اننا بحكمنا هذا نستطيع أن نقرر ما إذا كان افتراضنا الذي بدأنا به افتراض سليم أو غير سليم ، ذلك لأن الوضع الذي يمكن أن نصل اليه هو احد أمرين ، اما انه على أساس افتراضنا استنبطنا بيانات لا تختلف عن البيانات الواقعية الا اختلافاً بسيطاً وبذلك يكون هناك احتمال كبير في صحة الافتراض واما العكس إذا اختلفت البيانات المستنبطة عن البيانات الواقعية اختلافاً جوهرياً وبذلك يكون هناك احتمال كبير في عدم صحة الافتراض ، الأمر الذي يؤدي الى ضرورة رفضه واستبعاده .

ان بعض الأرقام النظرية يمكن ان تكون أقل بينا يمكن ان يكون بعضها أكثر من الأرقام الواقعية وبذلك فان جمع هذه الفروق بإشاراتها الجبرية يؤدي الى نتيجة مضللة في الحكم على المجموع الكلي للفروق بين التكرارات الواقعية والنظرية . ونظراً الى ان اهتمامنا ليس موجهاً الى كون الفرق بالزائد أو بالناقص وإنما لقيمة الفرق نفسه ، لذا وحق نتخلص من من الاشارات الجبرية ثربع الفروق . ولما كانت قيمة الفرق لا يمكن ان تعطينا فكرة سليمة من معنويته (اذا كان فرقاً بسيطاً أو كبيراً) حيث يمكن أن يكون فرقاً ما - واحد مثلاً - فرقاً كبيراً إذا كان نتيجة طرح أرقام صغيرة جداً ويمكن أن يكون فرقاً صغيراً جداً اذا كان نتيجة طرح أرقام

كبيرة جداً ، وبذلك حتى يمكن أن نأخذ فكرة سليمة عن معنوية الفرق ننسبه الى التكرارات النظرية التي حسبناها وفيما يلي هذه العمليات الحسابية منظمة في شكل جدول :

التكرارات الواقعية	التكرارات النظرية	الفرق	مربع الفرق	مربع الفرق التكرارات النظرية
٤٩٠	٥٤٠ر١	٥٠ر١	٢٥١٠ر٠١	٤ر٦٤٧٣
١٥٥٢	١٥٠١ر٩	٥٠ر١	٢٥١٠ر٠١	١ر٦٧١٢
٣٩٠	٣٣٩ر٩	٥٠ر١	٢٥١٠ر٠١	٧ر٣٨٤٦
٨٩٥	٩٤٥ر١	٥٠ر١	٢٥١٠ر٠١	٢ر٦٥٥٨
			المجموع	١٦ر٣٥٨٩

هذه القيمة ١٦ر٣٥٨٩ هي التي يطلق عليها اصطلاح كاي تربيع وهي التي تريد أن تحكم على معنويتها ، أي على ما اذا كانت قيمة جوهرية أو قيمة بسيطة يمكن أن تكون راجعة الى الصدفة . وحتى نستطيع ذلك يجب أن نرجع الى جدول كاي تربيع ، وهو جدول توزيع القيم المختلفة لكاي تربيع عند ما تكون الصدفة هي العامل الوحيد الذي يلعب دوره في إيجاد الفروق بين التكرارات المشاهدة (الواقعية) والتكرارات (النظرية) ، وقد اعد هذا الجدول كارل بيرسون ليكون اداة موضوعية (غير خاضعة للتقدير الشخصي للباحث) يمكن على أساسها الحكم على معنوية الفروق وتحديد احتمال الصدفة في الحصول عليها .

لا شك أن معنوية كاي تربيع (١٦ر٣٥٨٩ في مثالنا هذا) تتوقف على عدد القيم التي جمعت حيث يمكن أن تكون قيمة ذات معنى اذا كانت مجموعاً لعدد

بسيط جداً من القيم ويمكن أن تكون قيمة بسيطة اذا كانت مجموعاً لعدد كبير جداً من القيم . لهذا وحتى نستطيع أن نحكم على معنوية كاي تربيع حكماً دقيقاً يجب أن نحدد عدد القيم التي جمعت للتوصل الى هذه القيمة (كاي تربيع) . الا اننا نلاحظ اننا في تطبيقنا للافتراض الذي بدأنا به حتى نتوصل الى التكرارات النظرية لم نكن احرار مطلق الحرية في ذلك حيث كنا دائماً مقيدين بمجموع كل فئة من الفئات (مجموع المدخزين ومجموع غير المدخزين ثم مجموع العزاب ومجموع المتزوجين) ، اذ لا يجب أن نغير من هذه المجاميع عند تطبيق الافتراض ، وبمعنى آخر كانت حريتنا قاصرة على توزيع هذه المجاميع بين الفئات المختلفة ، لهذا كان يكفي أن نحسب التكرار النظري لاحدى هذه الفئات الاربع على أساس الافتراض اما باقي التكرارات النظرية فعند حسابها كنا مقيدين بالمجاميع الواقعية ، وبذلك فان معنوية كاي تربيع مرتبطة بدرجة حرية واحدة وليست بربع درجات ، وبمعنى آخر هناك تكرار واحد فقط الذي على أساسه يمكن أن نحكم على معنوية كاي تربيع وليس اربع تكرارات . هذا هو الأساس النظري لما نسميه بدرجات الحرية التي تظهر في كعب الجدول الذي أعده بيرسون فيما يختص بتوزيع كاي تربيع . ويمكن حساب عدد درجات الحرية حساباً سريعاً بطرح واحد من عدد الفئات الافقية التي يتكون منها الجدول وضرب الناتج في حاصل طرح واحد من عدد الفئات الرأسية التي تظهر في الجدول .

$$\text{أي } (1 - 2) (1 - 2) = 1 \text{ في مثالنا السابق .}$$

أو بطرح عدد الثوابت التي استخدمت في تقدير التكرارات النظرية من عدد المشاهدات (التكرارات النظرية) أي $1 = 3 - 2$

وبالرجوع الى جدول كاي تربيع وأمام درجة حرية واحدة نجد الوضع التالي :-

احتمال الحصول على قيم كاي تربيع بطريقة المصادفة							درجات الحرية
٠٠١ ر	٠١ ر	٠٥ ر	١٠ ر	٥٠ ر	٩٠ ر	٩٩ ر	
١٠٨٢٧ ر	٦٦٢٥ ر	٣٨٤١ ر	٧٠٦ ر	٤٥٥ ر	١٥٨ ر	١٥٧ ر	١

ولما كانت قيمة كاي تربيع في مثالنا السابق هي ١٦٣٥٨٩ فاحتمال الحصول على هذه القيمة بطريقة المصادفة احتمال صغير جداً أقل من ٠٠١ ر أي أقل من ١٪ .

والمشكلة التي تواجهنا الآن هي ما هو الاحتمال الذي يمكن أن نعتبره صغيراً وذلك الذي يمكن أن نعتبره احتمالاً كبيراً ؟ لقد جرى العرف في الدراسات الاجتماعية على ان احتمال الصدفة ٥٪ أو أقل يعتبر احتمالاً صغيراً، وانه إذا زاد احتمال الصدفة عن ٥٪ يعتبر احتمالاً كبيراً .

وبذلك فاحتمال ان نحصل على القيمة ١٦٣٥٨٩ بطريق المصادفة احتمال صغير جداً أي ان هذه القيمة تمثل فرقاً جوهرياً بين التكرارات الواقعية (المشاهدة) والتكرارات النظرية المتوقعة على أساس الافتراض الذي بدأنا به . وبمعنى آخر يمكن أن نقول انه على أساس افتراضنا بالتشابه بين سلوك العزاب والمتزوجين نحو الادخار أي انهم يكونوا مجتمعاً واحداً توقعنا بيانات اختلفت اختلافاً جوهرياً عن البيانات الواقعية وبهذا نستطيع أن نقول انه على أساس اختبارنا هذا هناك احتمال كبير في عدم صحة هذا الافتراض . لذلك فان القول بعدم وجود أي علاقة بين الحالة الزوجية وبين السلوك نحو الادخار قول يجب استبعاده حيث ان البيانات المتوفرة لدينا تقيم الدليل القوي على وجود العلاقة بين هاتين الظاهرتين ، وبمعنى آخر ندل البيانات المشاهدة (الواقعية) على ضعف الادخار بين العزاب

وشيوعه بين المتزوجين بدرجة تزيد عما نتوقع على أساس الافتراض الذي بدأنا به .

مثال ٣١ :

قام باحث في استهلاك نوع من الحلوى بتقسيم دولة ما إلى ثمان مناطق وأخذ من كل منطقة عينة عشوائية من السكان صنفهم تبعاً لاستهلاكهم وعدم استهلاكهم لهذا النوع ، وكانت البيانات التي حصل عليها كالآتي :

رقم المنطقة	عدد المستهلكين	عدد غير المستهلكين
١	٥٦	١٧
٢	٨٧	٢٠
٣	١٤٢	٥٨
٤	٧١	٢٠
٥	٧٧	٣١
٦	٧٢	٢٣
٧	١٠٠	٢٥
٨	١٤٢	٣٨
المجموع	٧٥٨	٢٤٢

ويريد الباحث أن يعرف ما إذا كان استهلاك هذا النوع من الحلوى يختلف اختلافاً جوهرياً بين الثمان مناطق التي تتكون منها الدولة أو ان الاختلاف ليس الا اختلافاً بسيطاً يمكن أن نعزوه الى عشوائية الدراسة بالمعينة .

وللاجابة على هـ الاستفسار نبدأ بفترض ان استهلاك هذا النوع من الحلوى لا يختلف بين هذه المناطق وانها جميعاً تكون مجتمعاً واحداً تكون فيه النسبة العامة لمستهلكي هذا النوع هي $\frac{758}{1000}$.

وبذلك فكل منطقة يمكن أن نعتبرها صورة صحيحة ينعكس عليها هذا السلوك العام نحو استهلاك هذا النوع من الحلوى - وعلى هذا الأساس يمكن أن نحسب عدد المستهلكين وغير المستهلكين في كل منطقة كالآتي (التكرارات النظرية) :

$$\text{المنطقة ١ - عدد المستهلكين } 73 \times \frac{758}{1000} = 55 \text{ تقريباً}$$

$$\text{وعدد غير المستهلكين } = 73 - 55 = 18$$

$$\text{المنطقة ٢ - عدد المستهلكين } 107 \times \frac{758}{1000} = 81 \text{ تقريباً}$$

$$\text{وعدد غير المستهلكين } = 107 - 81 = 26$$

$$\text{المنطقة ٣ - عدد المستهلكين } 200 \times \frac{758}{1000} = 152 \text{ تقريباً}$$

$$\text{وعدد غير المستهلكين } = 200 - 152 = 48$$

$$\text{المنطقة ٤ - عدد المستهلكين } 91 \times \frac{758}{1000} = 69 \text{ تقريباً}$$

$$\text{وعدد غير المستهلكين } = 91 - 69 = 22$$

$$\text{المنطقة ٥ - عدد المستهلكين } 119 \times \frac{758}{1000} = 90 \text{ تقريباً}$$

$$\text{وعدد غير المستهلكين } = 119 - 90 = 29$$

$$\text{المنطقة ٦ - عدد المستهلكين } ٩٥ = \frac{٧٥٨}{١٠٠٠} \times ٧٢ \text{ تقريباً}$$

$$\text{وعدد غير المستهلكين} = ٩٥ - ٧٢ = ٢٣ .$$

$$\text{المنطقة ٧ - عدد المستهلكين } ١٢٥ = \frac{٧٥٨}{١٠٠٠} \times ٩٥ \text{ تقريباً}$$

$$\text{وعدد غير المستهلكين} = ٩٥ - ١٢٥ = ٣٠ .$$

$$\text{المنطقة ٨ - عدد المستهلكين } ١٩٠ = \frac{٧٥٨}{١٠٠٠} \times ١٤٤ \text{ تقريباً}$$

$$\text{وعدد غير المستهلكين} = ١٩٠ - ١٤٤ = ٤٦ .$$

والخطوة التالية هي تحديد الفرق (كاي تربيع) بين التكرارات الواقعية والتكرارات النظرية :

المنطقة	التكرارات الواقعية (المشاهدة)	التكرارات النظرية (المتوقعة)	ف	ف ٢	ف ٢ التكرارات النظرية
١	٥٦	٥٥	١	١	١٨ ر ٠
	١٧]	١٨	١	١	٥٥ ر ٠
٢	٨٧	٨١	٦	٣٦	٤٤٥ ر
	٢٠]	٢٦	٦	٣٦	٣٨٤ ر ١
٣	١٤٢	١٥٢	١٠	١٠٠	٦٥٧ ر
	٥٨]	٤٨	١٠	١٠٠	٨٣ ر ٢
٤	٧١	٦٩	٢	٤	٥٧ ر ٠
	٢٠]	٢٢	٢	٤	١٨١ ر
٥	٨٨	٩٠	٢	٤	٤٥ ر ٠
	٣١]	٢٩	٢	٤	١٣٧ ر
٦	٧٢	٧٢	صفر	صفر	صفر
	٢٣]	٢٣	صفر	صفر	صفر
٧	١٠٠	٩٥	٥	٢٥	٢٦٣ ر
	٢٥]	٣٠	٥	٢٥	٨٣٤ ر
٨	١٤٢	١٤٤	٢	٤	٢٧ ر ٠
	٤٨]	٤٦	٢	٤	٨٧ ر ٠
المجموع					
					٦٢٧٣ ر

والخطوة التالية تحديد عدد درجات الحرية وهي في هذا المثال (٨ - ١) (٢ - ١) = ٧ ، وبشكل آخر ١٦ تكرار نظري ناقص عدد الثوابت التي استخدمت في حساب هذه التكرارات النظرية ، وهي في هذه الحالة ٩ ثوابت - أي ١٦ - ٩ = ٧ درجات حرية .

من جدول كاي تربيع نحصل على البيانات التالية :

احتمال الحصول على قيمة كاي تربيع بطريق المصادفة							درجات الحرية
٩٩ر	٩٠ر	٥٠ر	١٠ر	٥ر	١ر	٠٠١ر	٧
٢٣٩ر	٢٨٣٣ر	٦٣٤٦ر	١٢٠١٧ر	١٣٠٦٧ر	١٨٤٧٥ر	٢٤٣٣٢ر	

وبذلك فاحتمال الحصول على القيمة ٦٢٧٣ بطريق المصادفة اكثر من ٥٠٪ وهو احتمال كبير جداً ، وبذلك ليس هناك ما يجعلنا نشك في افتراضنا بان استهلاك هذا النوع من الحلوى متشابه بين المناطق المختلفة التي تتكون منها الدولة . وبشكل آخر يمكننا أن نقول ان هذا الاختبار لا يقيم الدليل ضد الفرض الذي وضعناه . وتفسير ذلك اننا على أساس افتراضنا بان البيانات المتوفرة لدينا (الواقعية) لا توحى بالاختلاف بين هذه المناطق بالنسبة الى استهلاك هذا النوع من الحلوى استنتجنا بيانات (التكرارات النظرية) لم تختلف عن البيانات الواقعية إلا اختلافاً بسيطاً احتمال أن يكون راجعاً إلى مجرد الصدفة احتمال كبير .

مثال ٣٢ :

قام و. ف. ولدون بتجربة رمي ١٢ زهرة نرد دفعة واحد وقد كرر هذه التجربة ٤٠٩٦ مرة ولاحظ في كل مرة عدد الزهرات الناجحة (النجاح في هذه التجربة هو ظهور ٦ أو ٥ أو ٤) فحصل على التوزيع الآتي :

التكرار المشاهد

عدد الزهرات الناجحة

صفر	صفر
٧	١
٦٠	٢
١٩٨	٣
٤٣٠	٤
٧٣١	٥
٩٤٨	٦
٨٤٧	٧
٥٣٦	٨
٢٥٧	٩
٧١	١٠
١١	١١
صفر	١٢

المجموع ٤٠٩٦

في هذا المثال نلاحظ ان احتمال النجاح بالنسبة للرهرة الواحدة =

$$\frac{3}{6} = \frac{1}{2} \text{ واحتمال الفشل } = \frac{3}{6} = \frac{1}{2} \text{ كذلك، ورمي ١٢ زهرة دفعة واحدة يعني عمل ١٢ محاولة في كل رمية . وعلى هذا الأساس يمكننا أن نحدد عدد المرات من (ال ٤٠٩٦ مرة) التي يمكن أن لا ينجح فيها أية زهرة وتلك التي يمكن أن ينجح فيها زهرة واحدة تم التي تنجح فيها زهرتين وهكذا .}$$

عدد المرات التي لا تنجح فيها زهرة واحدة =

$$١٢ \text{ ق. } ١ = ٤٠٩٦ \times \frac{١٢١}{٢} \times \frac{١}{٢} \times \frac{١}{٢} \text{ صفر ١ مرة}$$

عدد المرات التي تنجح فيها زهرة واحدة =

$$١٢ \text{ ق. } ١ = ٤٠٩٦ \times \frac{١١١}{٢} \times \frac{١}{٢} \times \frac{١}{٢} \text{ مرة}$$

عدد المرات التي تنجح فيها زهرتان =

$$١٢ \text{ ق. } ٢ = ٤٠٩٦ \times \frac{١٠١}{٢} \times \frac{٢١}{٢} \times \frac{١}{٢} \text{ مرة}$$

عدد المرات التي تنجح فيها ثلاث زهرات =

$$١٢ \text{ ق. } ٣ = ٤٠٩٦ \times \frac{٩١}{٢} \times \frac{٣١}{٢} \times \frac{١}{٢} \text{ مرة}$$

عدد المرات التي تنجح فيها أربع زهرات =

$$١٢ \text{ ق. } ٤ = ٤٠٩٦ \times \frac{٨١}{٢} \times \frac{٤١}{٢} \times \frac{١}{٢} \times \frac{١}{٢} \text{ مرة}$$

عدد المرات التي تنجح فيها خمس زهرات =

$$١٢ \text{ ق. } ٥ = ٤٠٩٦ \times \frac{٧١}{٢} \times \frac{٥١}{٢} \times \frac{١}{٢} \times \frac{١}{٢} \times \frac{١}{٢} \text{ مرة}$$

عدد المرات التي تنجح فيها ست زهرات =

$$١٢ \text{ ق. } ٦ = ٤٠٩٦ \times \frac{٦١}{٢} \times \frac{٦١}{٢} \times \frac{١}{٢} \times \frac{١}{٢} \times \frac{١}{٢} \times \frac{١}{٢} \text{ مرة}$$

عدد المرات التي تنجح فيها سبع زهرات =

$$١٢ ق \times \frac{١}{٦} \times \frac{١}{٦} \times ٤٠٩٦ = ٧٩٢ \text{ مرة}$$

عدد المرات التي تنجح فيها ثمان زهرات =

$$12 \text{ ق} = 1.96 \times \frac{1}{4} \times \frac{1}{4} \times 1 \text{ مرة}$$

عدد المرات التي تنجح فيها تسع زهرات =

$$12 \text{ ق} = 40.96 \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4}$$

عدد المرات التي تنجح فيها عشر زهرات =

$$١٢ ق. = \frac{101}{6} \times \frac{1}{6} \times ٤٠٩٦ = ٦٦ \text{ مرة}$$

عدد المرات التي تنجح فيها احد عشر زهرة =

$$12 \text{ مرة} = 4.96 \times \frac{1}{6} \times \frac{111}{6} \times 12 \text{ ق}$$

عدد المرات التي تنجح فيها اثني عشر زهرة =

$$١٢ ق ١٢ \times \frac{١}{٦} \times \frac{١٢٢}{٦} \times \frac{١}{٦} \times ٤٠٩٦ = ١ \text{ مرة}$$

يتبين لنا أن هنالا فرق بين عدد مرات النجاح المتوقعة وعدد مرات النجاح المشاهدة فعلا ونريد الآن أن نختبر الفرق بينها حتى نحدد ما إذا كان بسيطاً يمكن أن نزوه الى مجرد المصادفة أو أنه فرق جوهري يمكن على

أساسه أن شك في عدم تحيز زهر النرد الذي نقوم برميهِ في هذه التجربة .
وللاجابة على ذلك نقوم بإجراء اختبار كاي تربيع كالآتي :

عدد الزهرات الناجحة	التكرارات المشاهدة	التكرارات المتوقعة	ف	ف ^٢	ف ^٢ التكرار المتوقع
صفر	صفر	١	١ -	١	١٠٠٠٠
١	٧	١٢	٥ -	٢٥	٢٠٨٣٣
٢	٦٠	٦٦	٦ -	٣٦	٠٥٤٥٥
٣	١٩٨	٢٢٠	٢٢ -	٤٨٤	٢٢٢٠٠
٤	١٣٠	٤٩٥	٦٥ -	٤٢٢٥	٨٥٣٥٤
٥	٧٣١	٧٩٢	٦١ -	٣٧٢١	٤٦٩٨٢
٦	٩٤٨	٩٢٤	٢٤ +	٥٧٦	٠٦٣٣٤
٧	٨٤٧	٧٩٢	٥٥ +	٣٠٢٥	٣٨١٩٤
٨	٥٣٦	٤٩٥	٤١ +	١٦٨١	٣٣٩٦٠
٩	٢٥٧	٢٢٠	٣٧ +	١٣٦٩	٦٢٢٤٧
١٠	٧١	٦٦	٥ +	٢٥	٣٧٨٨
١١	١١	١٢	٢ -	٤	٣٠٧٧
١٢	صفر	١			
المجموع					
٣٣٨١٠٤					

وفي هذا المثال يكون عدد درجات الحرية (١٣ - ١) = ١٢ درجة

حيث طرحنا درجة واحدة نظراً لأننا كنا مقيدين بالمجموع الكلي لعدد المحاولات (٤٠٩٦ محاولة) عند حساب التكرارات المتوقعة . وبالنظر الى جدول كاي تربيع نحصل على البيانات التالية :

درجات الحرية	احتمال الحصول على قيمة كاي تربيع بطريقة المصادفة					
	٠٠١ر	٠١٠ر	٠٥ر	١٠ر	٥٠ر	٩٩ر
١٢	٣٢ر٩٠٩	٢٦ر٢١٧	١٢ر٠٢٦	١٨ر٥٤٩	١١ر٣٤٠	٦ر٣٠٤
						٣ر٥٧١

وبذلك يكون احتمال الحصول على القيمة ٣٣ر٨١٠٤ (قيمة كاي تربيع في مثالنا هذا) بطريق المصادفة احتمال ضعيف جداً ، فهي لذلك تشكل فرقاً جوهرياً بين التكرارات المشاهدة والتكرارات المتوقعة ، الأمر الذي يجعلنا نشك في عدم تحيز زهر النرد المستعمل في التجربة ، وبمعنى آخر فإن اختبار كان تربيع يوحي الينا بتحيز زهر النرد المستعمل في التجربة ، هذا اذا كان لدينا ثقة تامة في عشوائية القائم بالتجربة .

مثال ٣٣ :

في احدى التجارب التي قام بها مندل على استنبات بعض البذور حصل على ٣١٥ بذرة صفراء ومستديرة و ١٠١ بذرة صفراء ومجمدة و ١٠٨ بذرة خضراء ومستديرة و ٣٢ بذرة خضراء ومجمدة . هل تعتبر هذه النتيجة مطابقة للنظرية الوراثة الخاصة بهذا الموضوع حيث نتوقع ان توزع هذه البذور بين هذه الفئات الاربع بنسبة ٩ : ٣ : ١ .

على أساس هذا التوزيع المتوقع كان يجب ان نحصل على :

هذه القيمة فرقاً بسيطاً ، وبمعنى آخر فإن التكرارات المشاهدة لا تختلف عن التكرارات المتوقعة الا اختلافاً بسيطاً يمكن أن يكون راجعاً الى المصادفة.

مثال ٣٤ :

تدل البيانات التالية على عدد المواليد في إنجلترا وويلز عام ١٩٤١ مصنفة تبعاً لأشهر السنة . والمطلوب استخدام اختبار كاي تربيع لمعرفة ما اذا كانت هناك تغيرات موسمية في المواليد على أساس هذه البيانات :

الشهر	عدد المواليد	الشهر	عدد المواليد
يناير	٥٠١٥٩	يوليو	٤٩٣٩٥
فبراير	٤٥٨٨٥	اغسطس	٥٠٤٤٣
مارس	٥٠٨١٩	سبتمبر	٥١٥٦٢
ابريل	٤٩٠٧٠	اكتوبر	٥٠٢٢٤
مايو	٥٠٧٧١	نوفمبر	٤٧١٦٨
يونيو	٤٦٧٨٨	ديسمبر	٥٠٥٢٩
المجموع		٥٩٢٨١٣	

على أساس افتراض أن لا يوجد هناك تغيرات موسمية فان عدد المواليد اليومي يتوقع أن يكون $592813 \div 365 = 1624.1$ مولود تقريباً وبذلك يكون عدد المواليد لكل شهر من الأشهر ذات الثلاثين يوماً 48724 مولود ولكل شهر من الأشهر ذات الواحد والثلاثين يوماً 53049 مولود ولشهر فبراير 45476 مولود .

والخطوة التالية هي اختبار الفرق بين هذه البيانات المتوقعة والبيانات الواقعية كالآتي :

الشهر	عدد المواليد	العدد المتوقع	ف	ف ^٢	ف ^٢ العدد المتوقع
يناير	٥٠١٥٩	٥٠٣٤٩	١٩٠	٣٦١٠٠	٠٧١٧
فبراير	٤٥٨٨٥	٤٥٤٧٦	٤٠٩	١٦٧٢٨١	٣٦٨٧
مارس	٥٠٨١٩	٥٠٣٤٩	٤٧٠	٢٢٠٩٠٠	٤٣٨٧
ابريل	٤٩٠٧٠	٤٨٧٢٤	٣٤٦	١١٩٧١٦	٢٤٥٧
مايو	٥٠٧٧١	٥٠٣٤٩	٤٢٢	١٨٧٠٨٤	٣٥٣٧
يونيو	٤٦٧٨٨	٤٨٧٢٤	١٩٣٦	٣٧٤٨٠٩٦	٧٦٨٢٥
يوليو	٤٩٣٩٥	٥٠٣٣٩	٩٥٤	٩١٠١١٦	١٨٠٧٦
اغسطس	٥٠٤٤٣	٥٠٣٤٩	٩٤	٨٨٣٦	١٧٥
سبتمبر	٥١٥٦٢	٤٨٧٢٤	٢٨٣٨	٨٠٥٤٢٤٤	١٦٥٣٠٣
اكتوبر	٥٠٢٢٤	٥٠٣٤٩	١٢٥	١٥٦٢٥	٣١٠
نوفمبر	٤٧١٦٨	٤٨٧٢٤	١٥٥٦	٢٤٢١١٣٦	٤٩٦٩١
ديسمبر	٥٠٥٢٩	٥٠٣٤٩	١٨٠	٣٢٤٠٠	٦٤٣
المجموع	٥٩٢٨١٣	٥٩٢٨١٥			٣٢٥٨٩٩

وعلى أساس ١١ درجة حرية (١٢ - ١) فإن احتمال الحصول على القيمة ٣٢٥٨٩٩ بطريقة المصادفة احتمال ضعيف جداً وبذلك فإن هذا الاختبار يقيم الدليل على وجود تغيرات موسمية في المواليد حيث انه على اساس

افتراضنا بعدم وجود هذه التغيرات استنتجنا بيانات اختلفت عن الواقع
اختلافاً جوهرياً .

مثال ٣٥ :

البيانات الآتية تبين توزيع ٦٨٠٠ من الذكور في مدينة بادن تبعاً للون
الشعر ولون العين كذلك :

الجموع	لون الشعر				لون العين
	احمر	اسود	بني	اخر	
٣٨١١	٤٧	١٨٩	٨٠٧	١٧٦٨	ازرق
٣١٣٢	٥٣	٧٤٦	١٣٨٧	٩٤٦	رمادي او اخضر
٨٥٧	١٦	٢٨٨	٤٣٨	١١٥	بني او اسود
٦٨٠٠	١١٦	١٢٢٣	٢٦٣٢	٢٨٢٦	المجموع

والمطلوب دراسة ما اذا كانت هناك علاقة بين لون شعر الذكر ولون
عينيه تبعاً لهذه البيانات . وبمعنى آخر هل هناك علاقة بين تصنيف هذه
المجموعة من الذكور تبعاً للون الشعر ولون العينين او ان تصنيفهم تبعاً لهاتين
الظاهرتين مستقل عن الآخر .

للإجابة على ذلك نبدأ بافتراض عدم وجود اية علاقة بين لون الشعر ولون
العينين وان هذه المجموعة تكون مجتمعاً واحداً فيه نسبة ذوي العيون الزرق

$$\frac{2811}{6800} \text{ ونسبة ذوي العيون الخضراء او الرمادية } \frac{3132}{6800} \text{ ونسبة ذوي العيون السوداء}$$

$$\frac{857}{6800} \text{ وبتطبيق هذه النسب على المجموعات الاربع حسب لون الشعر نحصل}$$

على البيانات النظرية الآتية :

1169	$= \frac{2811}{6800} \times 2829$	مجموعة
ذوي عيون زرق		الرجال
1303	$= \frac{3132}{6800} \times 2829$	ذوي الشعر
ذوي عيون رمادية أو خضر		الأصفر
357	$= (1303 + 1169) - 2829$	
ذوي عيون بنية أو سود		

1088	$= \frac{2811}{6800} \times 2632$	مجموعة
ذوي عيون زرق		الرجال
1212	$= \frac{3132}{6800} \times 2632$	ذوي الشعر
ذوي عيون رمادية أو خضر		البنى
332	$= (1212 + 1088) - 2632$	
ذوي عيون بنية أو سود		

506	$= \frac{2811}{6800} \times 1223$	مجموعة
ذوي عيون زرق		الرجال
563	$= \frac{3132}{6800} \times 1223$	ذوي الشعر
ذوي عيون رمادية أو خضر		الأسود
154	$= (563 + 506) - 1223$	
ذوي عيون بنية أو سود		

$$\begin{aligned}
 & \text{مجموعة الرجال ذوي الشعر الأحمر} \\
 & ٦١٦ \times \frac{٧٨١١}{٦٨٠٠} = ٤٨ \text{ ذوي عيون زرق} \\
 & ٦١٦ \times \frac{٣١٣٢}{٦٨٠٠} = ٥٣٤ \text{ ذوي عيون رمادية او خضر} \\
 & ٦١٦ - (٥٣٤ + ٤٨) = ١٤٦ \text{ ذوي عيون بنية أو سود}
 \end{aligned}$$

بعد ذلك تختبر الفرق بين هذه البيانات المتوقعة (التكرارات النظرية) والبيانات الواقعية الواردة في الجدول السابق .

الصفة	التكرار الواقعي	التكرار النظري	ف	ف ^٢	ف ^٢ التكرار النظري
شعر اصفر و عيون زرق	١٧٦٨	١١٦٩	٥٩٩	٣٥٨٨٠١	٣٠٧
شعر اصفر و عيون خضر	٩٤٦	١٣٠٣	٣٥٧-	١٢٧٤٤٩	٩٨
شعر اصفر و عيون سود	١١٥	٣٥٧	٢٤٢-	٥٨٥٦٤	١٦٤
شعر بني و عيون زرق	٨٠٧	١٠٨٨	٢٨١-	٧٨٩٦١	٧٢
شعر بني و عيون خضر	١٣٨٧	١٢١٢	١٧٥	٣٠٦٢٥	٢٥
شعر بني و عيون سود	٤٣٨	٣٣٢	١٠٦	١١٢٣٦	٣٤
شعر اسود و عيون زرق	١٨٩	٥٠٦	٣١٧-	١٠٠٤٨٩	١٩٩
شعر اسود و عيون خضر	٧٤٦	٥٦٣	١٨٣	٣٣٤٨٩	٧٠
شعر اسود و عيون سود	٢٨٨	٥٤١	١٣٤	١٧٩٥٦	١١٧
شعر احمر و عيون زرق	٤٧	٤٨	١-	١	× ٠٠٠
شعر احمر و عيون خضر	٥٣	٥٣٤	٤-	١٦	٠٠٠
شعر احمر و عيون سود	١٦	١٤٦	٤	١٩٦	٠٠٠
× الأرقام ضئيلة ويمكن اهمالها					
١٠٧٦					

وعلى اساس درجات الحرية والتي تساوي في هذا المثال (٤-١) (٣-١)
 $= ٦$ او ١٢ تكرار نظري ناقصاً عدد الثوابت التي استخدمت في حساب
 هذه التكرارات وهي ٦ اي (١٢ - ٦) $= ٦$ درجات حرية وبالرجوع الى
 جدول كاي تربيع نجد ان احتمال الحصول على القيمة ١٠٧٦ بطريقة المصادفة
 احتمال ضعيف جداً ، وبذلك فهذا الاختبار يقيم الدليل على عدم صحة
 الافتراض بعدم وجود اية علاقة بين لون الشعر ولون العينين وبمعنى آخر بان هذا
 الاختبار يقيم الدليل على ان تصنيف هذه المجموعة من الذكور تبعاً للون
 الشعر ليس مستقلاً عن تصنيفهم تبعاً للون العينين .

ملاحظات على تطبيق اختبار كاي تربيع

١ - يدلنا اختبار كاي تربيع على احتمال الحصول في حالة المعاينة
 العشوائية على قيمة لكاي تربيع تساوي او اكثر من القيمة التي حصلنا عليها
 فعلاً بتطبيق الاختبار ، فاذا كان الاحتمال ضعيفاً يكون لنا الحق في الاعتقاد
 بوجود اختلاف جوهري بين البيانات المشاهدة المتوقعة . ولكنها لا نستطيع
 ان نسير في الاتجاه العكس فنقول - في حالة ان يكون الاحتمال ليس صغيراً -
 بصحة الافتراض الذي بدأنا به ، اذا ان كل ما نستطيع قوله ان الاختبار لا
 يقيم الدليل على عدم صحة الافتراض .

٢ - ان اختبار كاي تربيع ليس مقياساً لدرجة او نوع العلاقة بين
 الظواهر المختلفة . ان كل ما نستطيع ان نفهمه من هذا الاختبار هو ما اذا
 كان تصنيف مجموعة من الوحدات تبعاً لكل من الظاهرتين معينتين مستقل
 عن الآخر او غير مستقل عنه ولكننا لا نستطيع ان نفهم اي شيء فيما يختص
 بشكل العلاقة بينهما . ولتحديد درجة هذه العلاقة ونوعها لا بد من حساب
 مقاييس اخرى (مقاييس الارتباط) .

٣ - لا يمكن اجراء الاختبار ما لم تكن التكرارات في شكل مطلق

حيث انه اذا كانت التكرارات في شكل نسبي لا يمكن الحكم على مدى الاختلاف بين البيانات الواقعية والبيانات النظرية . والفرق بين القيمة المطلقة ٣ والقيمة ٤ يختلف كثيراً عن الفرق بين القيمة المطلقة ٣٠٠ والقيمة ٤٠٠ ، بينما اذا اعطينا هذه القيم في شكل نسبي فان الفرق في الحالتين يكون متساوي .

٤ - يجب ان يكون حجم العينة موضوع الدرس ليس صغيراً ، وقد اقترح يول وكندل (Yule and Kendall) الحجم ٥٠ كأقل حجم يمكن دراسته لاجراء الاختبار . وكذلك يجب ان لا يكون عدد الوحدات في كل فئة من الفئات صغيراً وقد جرى العرف على ان خمس وحدات هو اقل عدد يمكن ان يكون في اية فئة من الفئات ويحسن ان يكون عشرة . على ان هذه المشكلة يمكن التغلب عليها بتجميع الوحدات الموجودة في بعض الفئات اذا كان عددها صغيراً خاصة وان الفئات الصغيرة تكون دائماً في اطراف التوزيعات .

٥ - ان من محاسن اختبار كاي تربيع كأداة في الابحاث العلمية ان قيم كاي تربيع المحسوبة لعدة عينات خاصة بنفس الموضوع تحت الدرس ومأخوذة من نفس المجتمع يمكن اضافتها حتى يمكن ان يكون حكمنا افضل عما لو اخذنا كل عينة على حدة .

فاذا افترضنا اننا اخذنا أربع عينات من العمال الصناعيين وصنفنا عمال كل عينة على اساس توظيفهم وعلى أساس نوع الصناعة التي يعملون فيها فإننا نحصل على أربع مجموعات في كل عينة - العمال الموظفون والعاملون في الصناعات التي تنتج السلع الرأسمالية ، والعمال الموظفون في الصناعات التي تنتج السلع الاستهلاكية ، والعمال المتعطلون الذين يعملون عادة في الصناعات الرأسمالية ثم المتعطلون والذين يعملون عادة في الصناعات الاستهلاكية .

وعلى أساس البيانات الواقعية التي نحصل عليها وبافتراض عدم وجود أية علاقة بين التوظيف ونوع الصناعة سوف نستنتج بيانات نظرية ، ثم نقوم

باجراء اختبار كاي تربيع بالنسبة لكل عينة على حدة ونفترض اننا حصلنا على النتائج التالية :

رقم العينة	عدد درجات الحرية	كاي تربيع
١	١	٣٧٥
٢	١	٣٦٠
٣	١	٢١٢
٤	١	٤٢٠
المجموع	٤	١٣٦٧

من هذه النتائج ومن جدول كاي تربيع نجد ان العينات الثلاث الأولى تبين عدم وجود فرق جوهري بين البيانات الواقعية والبيانات النظرية باحتمال ٥ ٪ بينما تبين العينة الرابعة وجود فرق جوهري باحتمال ٥ ٪ ولكن ليس باحتمال ١ ٪ ، بينما على أساس مجموع كاي تربيع ومجموع درجات الحرية ٤ تبين وجود فرق جوهري باحتمال ٥ ٪ وكذلك باحتمال ١ ٪ . ولا شك اننا بذلك نستطيع أن نحكم حكماً يتفق مع ما نتوقعه حيث انه من المعتقد ان وقع البطالة يكون أكبر في الصناعات الرأسمالية بالمقارنة مع الصناعات الاستهلاكية . وتأکید وجود فرق جوهري تبعاً لمجموع كاي تربيع يؤيد ما نتوقعه ، حيث ان وجود فرق جوهري يقيم الدليل على عدم صحه افتراضنا بعدم وجود علاقة بين التوظيف ونوع الصناعة ولإدراك أهمية هذه الخاصية لاختبار كاي نورد المثال التالي :

يبين الجدول التالي نتيجة التلقيح ضد الكوليرا في احدى مناطق الدولة .

مثال ٣٦ :

المجموع	اصيب	لم يصب بالمرض	
٤٣٦	٥	٤٣١	لقح ضد الكوليرا
٣٠٠	٩	٢٩٠	لم يلحق ضد الكوليرا
٧٣٦	١٤	٧٢٢	المجموع

نفترض عدم وجود اية علاقة بين التلقيح ضد الكوليرا والاصابة بالمرض أي ان المجموعتين (الذين لقحوا والذين لم يلحقوا) يكونا مجتمعاً واحداً فيه نسبة الاصابة بالمرض $\frac{١٤}{٧٣٦}$ أي $\frac{٧}{٣٦٨}$. وبتطبيق هذه النسبة نحصل على البيانات النظرية التالية :

$$\left. \begin{array}{l} \text{الذين لقحوا} \\ \text{الذين لم يلحقوا} \end{array} \right\} \begin{array}{l} \text{يصاب بالمرض} \quad ٨٣ = \frac{٧}{٣٦٨} \times ٤٢٦ \\ \text{لا يصاب بالمرض} \quad ٤٢٧٧ = ٨٣ - ٤٣٦ \end{array}$$

$$\left. \begin{array}{l} \text{الذين لم يلحقوا} \\ \text{الذين لم يلحقوا} \end{array} \right\} \begin{array}{l} \text{يصاب بالمرض} \quad ٥٧ = \frac{٧}{٣٦٨} \times ٣٠٠ \\ \text{لا يصاب بالمرض} \quad ٢٩٤٣ = ٥٧ - ٣٠٠ \end{array}$$

ثم نجري الاختبار على الفرق بين البيانات الواقعية والبيانات النظرية كالآتي :

التكرار النظري	ف	ف	التكرار النظري	التكرار الواقعي	
٢٥ر٠	٨٩ر١٠	٣٣	٧٧ر٢٧٤	لم يصب ٤٣١	الذين لقحوا
١٢ر٣	٨٩ر١٠	٣٣	٨٣	اصيب ٥	
٣٧ر٠	٨٩ر١٠	٣٣	٣٢٩٤	لم يصب ٢٩١	الذين لم يلقحوا
١٠ر٩١	٨٩ر١٠	٣٣	٧٥	اصيب ٩	

وعلى أساس درجة حرية واحدة (٢ - ١) (٢ - ١) نلاحظ ان قيمة كاي تربيع ٣٢٨٤ واقعة بين احتمال ٥ ٪ واحتمال ١٠ ٪ وبذلك لا نستطيع أن نقرر قراراً فاصلاً في صحة أو عدم صحة افتراضنا .

لذلك يكون من الأفضل اجراء التجربة بالنسبة لعدة عينات من مناطق مختلفة ، فاذا فرضنا اننا حصلنا من هذه العينات النتائج التالية :

رقم العينة	عدد درجات الحرية	كاي تربيع
١	١	٢٨٣ر٣
٢	١	٤٣٩ر٩
٣	١	٨٠٦ر٦
٤	١	١٥٢ر٢
٥	١	١٦٥ر٥
٦	١	٩٥ر١
		١٤١ر٢٨

وعلى أساس ست درجات الحرية يكون احتمال الحصول على هذه القيمة :
٢٨٤١ بطريق المصادفة احتمال ضعيف جداً (أقل من ١ في الالف) ،
وبذلك فهذه القيمة تشكل فرقاً جوهرياً بين البيانات الواقعية والبيانات
النظرية ، الأمر الذي يقيم الدليل على عدم صحة الافتراض بعدم وجود علاقة
بين التطعيم ضد الكوليرا والاصابة بهذا المرض

اختبار حسن التطابق لتوفيق المنحنى المعتدل :

قمنا في الفصل السابق بتوفيق المنحنى المعتدل (بطريقة المساحات)
وتوصلنا بذلك إلى تكرارات تختلف بعض الشيء عن التكرارات الواقعية ،
وأشرنا إلى ان الخطوة التالية تكون في اختبار حسن التطابق بين هذه
التكرارات النظرية والتكرارات المشاهدة الواقعية . والجدول التالي يوضح
كيفية اجراء هذا الاختبار :

مثال :

التكرار الواقعي	التكرار النظري	ف	ف ^٢	ف ^٢	التكرار النظري
٦	٦	صفر	صفر	صفر	صفر
١٧	١٢	٥	٢٥	٢٠٨	٢٠٨
٢٥	٢٦	٩	٨١	٣١١	٣١١
٤٨	٤٨	صفر	صفر	صفر	صفر
٦٥	٧٩	١٤	١٩٦	٢٤٨	٢٤٨
٩٠	١١٣	٢٣	٥٢٩	٤٦٨	٤٦٨
١٣١	١٤٠,٥	٩,٥	٩٠,٢٥	٦٤	٦٤
١٧٣	١٥١	٢٢	٤٨٤	٣٠١	٣٠١
١٥٥	١٤٠,٥	١٤,٥	٢١٠,٢٥	١٥٠	١٥٠
١١٧	١١٣	٤	١٦	١٤	١٤
٧٥	٧٩	٤	١٦	٢٠	٢٠
٥٢	٤٨	٤	١٦	٣٤	٣٤
٢١	٢٦	٥	٢٥	٩٥	٩٥
٩	١٢	٣	٩	٧٥	٧٥
٦	٦	صفر	صفر	صفر	صفر
				١٩٨٨	١٩٨٨

وعلى أساس ١٢ درجة حرية (١٥ تكرار نظري - عدد الثوابت التي استخدمت في حساب هذه التكرارات وهي مجموع التكرارات والانحراف

المعياري والوسط الحسابي للتوزيع الواقعي) يتبين من جدول كاي تربيع ان احتمال الحصول على القيمة ١٨٩٣ أقل من ٥ ٪ أي احتمال ضعيف ، وبذلك لا يمكن أن نمنزو هذا الفرق إلى المصادفة ، فهو اذن فرق جوهري على أساسه يتعين علينا استبعاد الافتراض الذي بدأنا به وهو تماثل التوزيع الخاص بأجور العمال موضوع الدرس .

وبالمثل يمكن اختبار حسن التطابق بين التكرارات الواقعية والتكرارات النظرية التي توصلنا إليها بطريقة الاحداثيات الرأسية .

تمارين

١ - بين الجدول الآتي توزيع الطلبة الناجحين والراسبين :

صف الاستاذ أ	صف الاستاذ ب	صف الاستاذ ج
الناجحون ٥٠	٤٧	٥٦
الراسبون ٥	١٤	٨

اختبر معنوية الفرق بين نسب النجاح عند الاساتذة الثلاث :

(على أساس مستوى معنوية ٥ ٪ لا يمكن رفض الفرضية بأن لا فرق جوهري بين نسب النجاح عند الاساتذة الثلاث) .

٢ - بين الجدول التالي عدد الكتب المعارة في أيام الاسبوع لزائري إحدى المكتبات العامة :

الأيام :	الاثنين	الثلاثاء	الاربعاء	الخميس	الجمعة
عدد الكتب :	١٣٥	١٠٨	١٢٠	١١٤	١٤٦

هل هناك فرق جوهري بين عدد الكتب المعارة في هذه الايام ؟

(على اساس مستوى معنوية ٥ ٪ لا يمكن رفض الفرضية بان لا فرق جوهري بين عدد الكتب المعارة في هذه الايام) .

٣ - الآتي توزيع الطلبة تبعاً لدرجاتهم في الرياضيات وفي الطبيعة :

عالية في الطبيعة	متوسطة في الطبيعة	ضعيفة في الطبيعة	
٥٦	٧١	١٢	عالية في الرياضيات
٤٧	١٦٣	٣٨	متوسطة » » »
١٤	٤٢	٨٥	ضعيفة » » »

اختبر الفرضية بان درجات الرياضيات مستقلة عن درجات الطبيعة .
(على اساس مستوى معنوية ٥ ٪ يمكن رفض الفرضية) .

الفصل العاشر

العينات

سبق ان ذكرنا أن الدراسات الاحصائية يمكن أن تجري بالعينه لاستنتاج المقاييس التي نرغب في الوصول اليها والتي تمثل المجتمع الذي سحبت منه العينة . كذلك أشرنا الى أن المقاييس المستنتجة من العينة لا يمكن أن تكون هي نفسها المقاييس الخاصة بالمجتمع اذ لا بد أن يوجد فرق بينها ، هذا الفرق سميناه بخطأ العينة . فلو أجرينا دراسة عن أطوال الطلبة بالعينة وتبين منها أن متوسط طول الطالب ١٧٢ سم مثلاً فهل معنى ذلك أن متوسط طول الفرد في المجتمع يساوي فعلاً ١٧٢ سم ، وبمعنى آخر هل اذا قمنا بدراسة مجتمع الطلبة بالعد الشامل هل نصل الى هذه النتيجة فعلاً ؟ كذلك اذا قمنا بدراسة نسبة البطالة في المجتمع عن طريق عينة وتبين منها أن هذه النسبة هي ١٥٪ فهل يعني ذلك أن نسبة البطالة في المجتمع هي فعلاً ١٥٪ ؟ .

كذلك بينا فيما سبق أن نتائج العينات يمكن أن تتعرض لأخطاء التحيز ، على أن مثل هذه الأخطاء يجب القضاء عليها بسحب العينة سحباً علمياً سليماً ، ولكن بالرغم من ذلك قد لا يتحقق التطابق بين نتائج العينة ونتائج المجتمع الذي سحبت منه لو اجريت الدراسة بالعد الشامل . لذلك يكون من الواجب أن نتعرف على القوانين والعوامل التي تتحكم في الفرق بين المقاييس المستنتجة من العينة ومقاييس المجتمع الفعلية ، أي التي نتحكم في أخطاء

المعاينة . وحق نستطيع أن نتعرف على هذه القوانين والعوامل يجب أن ندرس التوزيع الاحتمالي لنتائج العينات .

نفترض أن لدينا مجتمعاً معيناً عدد وحداته N واردنا أن نسحب منه العينات الممكنة والتي يكون حجم كل منها n فما هو عدد هذه العينات ؟ وبمعنى آخر ما عدد الطرق التي يمكن بها اختبار عينة عدد وحداتها n من مجموع عدد وحدات المجتمع N ؟ ان عدد العينات التي يمكن

سحبها هو n أي $\frac{N!}{n!(N-n)!}$ وللتبسيط نفترض اننا نستطيع

أن نسحب ١٠٠٠ عينة من هذا المجتمع على أساس هذا الحجم للعينة . نقوم بعد ذلك بحساب المتوسط لكل عينة وبذلك يتكون لدينا ١٠٠٠ وسط يمثل جميع العينات التي أمكن سحبها من هذا المجتمع . نوزع هذه المتوسطات في توزيع تكراري ، وبذلك نحصل على عدد العينات التي يكون متوسطها واقعاً في كل فئة من فئات التوزيع . هذا التوزيع نسميه بتوزيع المعاينة ويمكن تعريفه بأنه توزيع متوسطات جميع العينات التي يمكن سحبها من مجتمع معين على أساس حجم معين لكل عينة من هذه العينات . ويتميز هذا التوزيع بالصفات الآتية :

١ - انه يكون توزيعاً معتدلاً أو قريباً جداً من الاعتدال .

٢ - ان الوسط الحسابي لهذا التوزيع يكون هو نفسه المتوسط الحسابي للمجتمع الذي سحبت منه العينات أو قريباً جداً منه . وبذلك نستطيع أن نقول ان متوسط متوسطات جميع العينات التي يمكن سحبها من مجتمع معين يكون هو نفسه المتوسط الواقعي للمجتمع الذي نسحب منه هذه العينات .

٣ - حيث ان توزيع المعاينة يكون توزيعاً معتدلاً بذلك نستطيع أن نطبق خاصية هذا التوزيع فنقول ان ٦٨ ٪ من العينات التي يمكن سحبها من

مجتمع معين لا يمكن أن يزيد متوسطها أو يقل عن متوسط المجتمع الأصلي إلا بمقدار ١ ع ، وان ٩٥٤٪ من العينات التي يمكن سحبها من مجتمع معين لا يمكن أن يزيد متوسطها أو يقل عن الوسط الواقعي للمجتمع إلا بمقدار ٢ ع ، وان ٩٩٧٪ من العينات التي يمكن سحبها من مجتمع معين لا يمكن أن يزيد متوسطها أو يقل عن الوسط الواقعي إلا بمقدار ٣ ع .

وبشكل آخر نستطيع أن نقول ان احتمال أن يكون المتوسط المقدّر من عينة أكبر أو أقل من المتوسط الواقعي للمجتمع بمقدار ١ ع هو ٦٨٪ . وان احتمال أن يكون المتوسط المقدّر من عينة أكبر أو أقل من المتوسط الواقعي للمجتمع بمقدار ٢ ع هو ٩٥٤٪ . وان احتمال ان يكون المتوسط المقدّر من عينة أكبر أو أقل من المتوسط الواقعي للمجتمع بمقدار ٣ ع هو ٩٩٧٪

٤ - ان الانحراف المعياري الذي تكلمنا عنه في الفقرة السابقة هو الانحراف المعياري لتوزيع المعاينة ، ونلاحظ ان هذا التوزيع هو توزيع نظري فقط أي لا يوجد لدينا عملياً حيث انه يمثل متوسطات جميع العينات التي يمكن سحبها من المجتمع ونحن في الواقع لا نسحب إلا عينة واحدة من هذا المجتمع . وبناء على الصفة السابقة لتوزيع المعاينة نستطيع أن نقول ان هذه العينة التي سحبناها تقع ضمن هذا التوزيع النظري ونستطيع أن نحدد مكان وقوعها فيه باحتمالات مختلفة ، فباحتمال ٦٨٪ تقع هذه العينة في القطاع الاول (س - + ١ ع ، س - - ١ ع) وباحتمال ٩٥٤٪ تقع في القطاع الثاني (س - + ٢ ع ، س - - ٢ ع) وباحتمال ٩٩٧٪ تقع في القطاع الثالث (س - + ٣ ع ، س - - ٣ ع) .

ان الانحراف المعياري لتوزيع المعاينة يتوقف اولاً على الانحراف المعياري للمجتمع الأصلي الذي نسحب منه العينات وعلى حجم العينات التي نسحبها والرسم التالي يوضح لنا كيف ان الانحراف المعياري لتوزيع المعاينة يقل

تدريجياً كلما زاد عدد الوحدات التي نسحبها في كل عينة . أي ان الانحراف المعياري لتوزيع المعاينة يتناسب طردياً مع الانحراف المعياري للمجتمع وعكسياً مع حجم العينات التي تسحب منه . وبالبحث الرياضي وجد ان الانحراف

المعياري لتوزيع المعاينة $= \frac{\sigma}{\sqrt{n}}$ ، على أن القانون الأصلي

للالانحراف المعياري لتوزيع المعاينة هو $= \frac{\sigma}{\sqrt{n}} \sqrt{\frac{n-1}{n}}$

ولكن عندما يكون حجم المجتمع لانهاياً تقترب قيمته $\sqrt{\frac{n-1}{n}}$

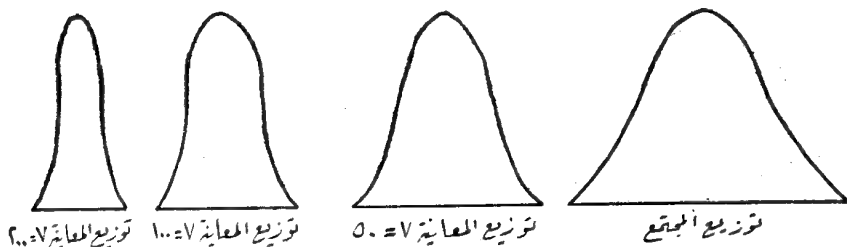
من الواحد الصحيح وبذلك يكون الانحراف المعياري لتوزيع المعاينة =

$\frac{\sigma}{\sqrt{n}}$ وهو القانون الشائع الاستعمال في الدراسات الاحصائية . ويسمى

الانحراف المعياري لتوزيع المعاينة تمييزاً له عن الانحراف المعياري للمجتمع بالخطأ المعياري . ويقاس لنا هذا المقياس مدى التشتت بين متوسطات العينات نتيجة عوامل الصدفة في سحب هذه العينات ، أي انه بالتالي يقيس الاخطاء العشوائية المتوقعة التي يتعرض له المتوسط المقدر من عينة . وفي الحالات التي لا يتوفر فيها الانحراف المعياري للمجتمع الذي سحبت منه العينة يمكن ان نستعاض بدلا عنه بالانحراف المعياري لوحدات العينة نفسها ، ومن الواضح ان هذا الانحراف المعياري لا يمكن ان يكون هو نفسه الانحراف المعياري للمجتمع ، الا أن الفرق بينهما يكون ضئيلاً إذا أمكن سحب العينة بحيث تمثل الاختلافات بين وحدات المجتمع تمثيلاً صحيحاً .

وخلاصة ما تقدم نقول انه عندما نحسب الخطأ المعياري للعينة يكون

ذلك باحتمال معين أي بدرجة ثقة ليست كاملة ، كما أن هذا الخطأ يكون احتمال الزيادة عن الوسط الواقعي للمجتمع واحتمال النقص عنه . وقد يختلف



الاحصائيون في تحديد حدود الثقة ، الا اننا سنعتبر ان وقوع وسط العينة (أو قيمة وحدة معينة في حالة التوزيع المعتدل لمجتمع احصائي) في حدود درجة الثقة 90% يجعلنا نستنتج انه ليس هناك ما يدعو الى الشك في عشوائية العينة أي أن الفرق بين متوسطها ومتوسط المجتمع هو فرق راجع الى الصدفة أي أنه غير معنوي أو غير جوهري . واذا وقع متوسط العينة بين درجتي الثقة 90% و 99% يكون هناك شك في عشوائية العينة . واذا وقع متوسط العينة خارج درجة الثقة 99% نؤكد عدم عشوائية العينة أي أن الفرق بينها وبين المجتمع فرق جوهري أي معنوي أي لا يمكن أن يكون راجعاً لمجرد الصدفة. وعلى أساس درجة الثقة 99% يكون حدي الثقة هما س- + 2ر58ع ، س- - 2ر58ع . أما اذا اعتبرنا درجة الثقة 90% يكون حدي الثقة هما : س- - 1ر96ع ، س- + 1ر96ع

استخدام متوسط العينة في تقدير متوسط المجتمع :

مثال ٣٨ :

نفترض اننا أخذنا عينة من الأطفال مكونة من ١٠٠ طفل تبين منها ان

متوسط الوزن هو ٣٠.٥ كيلو بانحراف معياري = ٩ كيلو ونريد أن نجد حدود متوسط وزن الاطفال من نفس العمر في المجتمع .

$$r = \frac{9}{\sqrt{100}} = \text{الخطأ المعياري للعينة}$$

وبذلك بدرجة ثقة ٩٥٪ يكون متوسط وزن الأطفال في المجتمع واقعاً في فترة الثقة :

$$30.5 + 9 \times 1.96$$

$$30.5 + 17.64 \quad \text{أي بين}$$

$$32.264 \text{ و } 28.736 \text{ كيلو أي بين}$$

وبدرجة ثقة ٩٩٪ يكون متوسط وزن الأطفال في المجتمع واقعاً في فترة الثقة :

$$30.5 + 9 \times 2.58$$

$$30.5 + 23.22 \quad \text{أي بين}$$

$$32.822 \text{ و } 28.178 \text{ كيلو أي بين}$$

ونلاحظ أن المدى في الحالة الثانية أوسع منه في الحالة الاولى وذلك طبيعي حيث أننا في الحالة الثانية نكون أكثر وثوقاً من النتيجة التي نصل إليها بالمقارنة مع نتيجة الحالة الأولى .

اختبار عشوائية العينة أي اختبار معنوية الفرق بين متوسط العينة

ومتوسط المجتمع :

إذا توفر لنا متوسط المجتمع الذي سحبت منه عينة ما فإننا نستطيع بمقارنة هذا المتوسط بالمتوسط المقدّر من العينة على أساس الدرجات المعيارية أن نختبر عشوائية العينة أي مدى تمثيلها للمجتمع وبمعنى آخر نستطيع أن نحكم على معنوية الفرق بينها وبين المجتمع ، ويكون لذلك أهمية كبيرة حيث إذا تبيننا عدم وجود فرق معنوي بين العينة والمجتمع الذي سحبت منه فإننا نطمئن إليها ونستمر في دراستها .

مثال ٣٩ :

إذا كان متوسط الأجر في صناعة ما هو ٩٦ ليرة في الأسبوع وأخذت عينة من المشتغلين في هذه الصناعة وكان توزيعهم حسب فئات الأجر كالآتي :

فئات الأجر	عدد العمال	م	ح	ح ك	ح ^٢ ك
٧٠ -	٧	٧٥	٢ -	٢٤ -	٢٨
٨٠ -	١٣	٨٥	١ -	١٢ -	١٣
٩٠ -	٢٥	٩٥	صفر	صفر	
١٠٠ -	١٨	١٠٥	١ +	١٨ +	١٨
١١٠ -	١٢	١١٥	٢ +	٢٤ +	١٨
١٢٠ - ١٣٠	٦	١٢٥	٣٩	١٨ +	٥٤
المجموع	٨١	-	-	٣٣	١٦١

$$\text{س} = ٩٥ + \frac{٣٣}{٨١} \times ١٠ = ٩٩.٠٧ \text{ ليرة}$$

$$ع = 10 \sqrt{\frac{161}{81} \left(\frac{33}{81}\right)^2} = 13.49 \text{ ليرة}$$

∴ الانحراف المعياري لتوزيع المعاينة (الخطأ المعياري) =

$$13.49 = \frac{13.49}{\sqrt{81}}$$

$$∴ \text{ الدرجة المعيارية} = \frac{96 - 99.07}{13.49} = 2.05$$

∴ متوسط العينة واقع بين الدرجة المعيارية 1.96 والدرجة المعيارية 2.05 . وبذلك يكون هناك شك في عشوائية العينة .

ويمكن أن نصل الى نفس النتيجة بطريقة اخرى إذ نستطيع أن نحسب فترة الثقة بدرجة 95 % :

$$99.07 + \underline{13.49 \times 1.96}$$

$$\underline{99.07 + 26.361}$$

$$\text{اي بين } 102.0061 , 125.4339$$

وحيث أن المتوسط الواقعي لا يقع في هذا المدى فلا يمكن التأكد من عشوائية العينة . وبدرجة ثقة 99 % يكون المدى :

$$99.07 + \underline{13.49 \times 2.058}$$

$$\underline{99.07 + 27.648} \text{ أي بين}$$

$$126.718 , 126.718$$

وحيث أن المتوسط الواقعي (٩٦) يقع في هذا المدى فيكون بذلك محصوراً بين درجة الثقة ٩٥٪ ودرجة الثقة ٩٩٪ ، أي المدى الذي يجعلنا نشك في عشوائيه العينة .

وبنفس الطريقة السابقة يمكن اختيار معنوية الفرق بين متوسط مجتمع معين ومتوسط جزء من هذا المجتمع على افتراض أن هذا الجزء يمثل عينة عشوائية من المجتمع لا تختلف في صفاتها عن صفات المجتمع كله . فإذا تبين لنا عدم وجود فرق معنوي نستطيع أن نصل الى النتيجة بأن هذا الجزء يمكن ان يعتبر عينة ممثلة تمثيلاً صادقاً للمجتمع ، أما اذا تبين لنا وجود فرق جوهري نستطيع أن نصل الى النتيجة بأن هناك عوامل معينة تجعل هذا الجزء يختلف اختلافاً جوهرياً عن المجتمع وبالتالي لا يمكن اعتباره عينة عشوائية سليمة من هذا المجتمع ، ويمكننا أن نستفيد من ذلك في مقارنة صفات المجتمعات الاحصائية التي ندرسها بصفات أجزائها ، فمثلاً نستطيع أن نقارن بين متوسط انتاج مصنع معين وانتاج جميع المصانع الشبيهة ، وبين انتاج منطقة معينة في الدولة من محصول معين وانتاج الدولة بأكملها من هذا المحصول ، وبين متوسط وزن الاطفال في مدرسة معينة وبين متوسط وزن الاطفال جميعاً في نفس العمر ... الخ

مثال ٤٠ :

كان متوسط محصول الفدان الواحد من القطن في احدى المحافظات عام ١٩٦٢ = ٤٢٠ رطلاً وكان الانحراف المعياري لغلة الفدان ٩٤ رطلاً ، فإذا علمنا أن المتوسط العام لغلة الفدان في الدولة هو ٤٣١ رطلاً ، اختبر معنوية الفرق بين المتوسطين اذا كانت المساحة المزروعة قطناً في هذه المحافظة هي ٢٠٠ فدان .

$$\text{الخطأ المعياري} = \frac{٩٤}{\sqrt{٢٠٠}} = ٦.٧ \text{ رطل} .$$

$$\text{الدرجة المعيارية} = \frac{٤٢٠ - ٤٣١}{٦.٧} = \frac{١١}{٦.٧} = ١.٦ \text{ رطل} .$$

وبذلك يكون متوسط المحافظة داخل درجة ثقة ٩٥ ٪ ، وبذلك لا يكون هناك ما يدعو إلى رفض الفرض القائل بأن أرض هذه المحافظة المزروعة قطعاً تعتبر عينة عشوائية من أراضي الدولة المزروعة بهذا المحصول .

وبطريقة أخرى يكون أقصى فرق بين المتوسطين بدرجة ثقة ٩٥ ٪ :

$$= ٦.٧ \times ١.٩٦ = ١٣.١٣٢ \text{ رطل} .$$

وحيث ان الفرق بين المحافظة والمجتمع هو ١١ رطل فقط أي انه أقل من ١.٩٦ خطأ معياري ، بذلك يكون الفرق في حدود ما تأتي به الصدفة العشوائية .

وفيدنا الاختبار السابق في الاجابة على السؤال - هل لعامل معين أثر على جزء من المجتمع - أي هل له تأثيره في اختلاف متوسط الجزء عن متوسط المجتمع؟ ولا شك انه لا يجب ان نتسرع ونحكم بذلك فقد يكون الفرق راجعاً لمجرد الصدفة العشوائية . ولذلك على اساس هذا الاختبار نستطيع أن نحكم في الموضوع . فاذا لم يكن هناك أي شك في عشوائية العينة أي إذا كان متوسطها لا يختلف عن الوسط العام للمجتمع زيادة او نقصاً إلا بمقدار ١.٩٦ خطأ معياري في الحد الأقصى فلا يكون لدينا أي شيء يدعو إلى الاعتقاد بوجود تأثير لهذا العامل . أما إذا كان هناك شك في عشوائية العينة ، أي إذا كان متوسطها بين درجة ثقة ٩٥ ٪ و ٩٩ ٪ يكون هناك ترجيح لوجود

تأثير لهذا العامل . وإذا كان في حكم المؤكد أن العينة غير عشوائية أي إذا خرج متوسطها عن درجة الثقة ٩٩٪ يكون في حكم المؤكد وجود تأثير لهذا العامل ، على انه يجب أن نلاحظ عدم وجود أي عامل آخر يمكن أن يكون له تأثير أي لا بد أن تتشابه الظروف بين الجزء والمجتمع ولا يكون هناك من فرق بينهما إلا بالنسبة لهذا العامل فقط وهو الذي ندرس تأثيره .

مثال ٤١ :

إذا علمنا ان متوسط انتاج العامل الواحد في صناعة ما هو ٦٠ قطعة بانحراف معياري ٧ قطع - وإذا علمنا أن عينة من ٩٠ عامل دربوا على العمل بطريقة أفضل فأصبح متوسط انتاجهم ٦٥ قطعة ، فهل نستنتج من ذلك أن هذا التدريب كان له أثره على تحسين الانتاج :

$$\text{بدرجة ثقة } ٩٥\% \text{ يكون المدى بين } ٦٠ \pm \frac{٧}{\sqrt{٩٠}} \times ١.٩٦$$

أي بين ٥٨.٥٥ ، ٦١.٤٥

$$\text{وبدرجة ثقة } ٩٩\% \text{ يكون المدى بين } ٦٠ \pm \frac{٧}{\sqrt{٩٠}} \times ٢.٥٨$$

أي بين ٥٨.١ ، ٦١.٩

وحيث ان متوسط الانتاج بعد التدريب أصبح ٦٥ قطعة ، إذا نستطيع أن نستنتج أن التدريب يؤدي فعلاً إلى تحسين الانتاج .

اختبار الفرق بين متوسطي عينتين :

قد يتوفر لدينا بيانات عن مجموعتين متماثلتين في جميع الظروف فيما عدا

عامل واحد يمكن أن يكون قد اثر في احدى المجموعتين دون الاخرى ، او قد اثر فيها بدرجات متفاوتة ، ونريد أن نتثبت فعلاً من تأثير هذا العامل ، وفي هذه الحالة لا يجب أن تتسرع ونقول أن الفرق بين المجموعتين أو العينين راجع فعلاً إلى هذا العامل ، إذ يجب علينا أولاً أن نتحقق من امكان أن يكون الفرق بينهما راجع إلى مجرد الصدفة ، فإذا استطعنا أن نستبعد هذا الفرض (فرض الصدفة) أمكننا أن نؤكد تأثير هذا العامل .

لنفترض مثلاً ان إحدى المؤسسات قامت بتطبيق طريقتين من طرق الانتاج وتبين من اتباعها أن متوسط جودة الانتاج قد اختلف ، فهل يكون هذا الاختلاف راجع إلى الصدفة أو هو اختلاف جوهري يجعلنا نؤكد تفضيل إحدى الطريقتين على الاخرى ؟ كذلك إذا أخذنا عينتين من الأراضي الزراعية واستعملنا في إحداها نوعاً من السماد وفي الاخرى نوعاً آخر أو تركناها بدون تسميد وفي نهاية الموسم لاحظنا ان متوسط انتاج العينة الاولى اختلف عن متوسط انتاج العينة الثانية فهل يكون الفرق راجعاً إلى الصدفة أو هو فرق جوهري ، الأمر الذي يجعلنا نؤكد تفضيل نوع على آخر أو تأثير السماد على الانتاج إذا لم تكن الأرض في العينة الاخرى قد سمدت .

والسؤال الذي نريد الاجابة عليه في مثل هذه الحالة ، هو هل يمكن اعتبار المجموعتين أو العينتين منتميتان إلى مجتمعين لهما نفس الوسط الحسابي أو لا ؟ وبمعنى آخر ان اختلاف هذا العامل المؤثر لا يقضي على احتمال كون العينتين من مجتمعين لهما نفس الوسط الحسابي ، وبذلك يكون الاختلاف بينهما راجعاً إلى الصدفة ، أو انها ينتميان إلى مجتمعين مختلفين تماماً ، وبذلك نؤكد تأثير هذا العامل .

إذا رمزنا إلى متوسط العينة الاولى بالرمز س-١ ، ومتوسط العينة الثانية بالرمز س-٢ ، والفرق بينهما بالرمز ف ، يكون هذا الفرق هو إحدى الفروق التي يتكون منها توزيع لجميع المروق المحتملة بين العينات التي يمكن اخذها

من المجتمعين الأصليين والتي يكون حجم كل عينة من المجتمع الاول n_1 ومن المجتمع الثاني n_2 . وفي حالة تساوي الوسطين الحسابيين للمجتمعين فان الوسط الحسابي للفروق المحتملة بينها لا بد ان يساوي صفراً حيث ان كل فرق موجب لا بد ان يقابله فرق آخر سالب . ويكون توزيع هذه الفروق توزيعاً معتدلاً وسطه الحسابي يساوي صفراً وتباينه يساوي مجموع تباين الوسطين الحسابيين s_1^2 ، s_2^2 .

وحيث ان تباين المجتمع الأول للعينات $= \frac{s_1^2}{n_1}$ وتباين المجتمع الثاني

$$\text{للعينات} = \frac{s_2^2}{n_2} \text{ يكون تباين توزيع الفروق بينها} = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$$

$$\text{أي ان } F^2 = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$$

$$\text{ويكون بذلك الخطأ المعياري لتوزيع الفروقي} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

وحيث ان توزيع مجتمع الفروق توزيع معتدل وسطه الحسابي = صفر ، نستطيع ان نطبق قواعد التوزيع المعتدل فنقول ان :

$$\text{بدرجة ثقة } 95\% \text{ تكون فترة الثقة} = \text{صفر} \pm 1.96$$

$$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\text{وبدرجة ثقة } 99\% \text{ تكون فترة الثقة} = \text{صفر} \pm 2.58$$

$$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

وبذلك إذا وقع الفرق بين العيين داخل فترة الثقة الاولى لا يكون هناك ما يبعث على الشك في ان تكون العينتان مأخوذتان من مجتمعين لهما نفس الوسط الحسابي أي ان الفرق بينهما غير جوهري ويمكن أن يكون راجعاً إلى الصدفة . أما إذا وقع الفرق بين العينتين بين فترة الثقة الاولى وفترة الثقة الثانية يكون هناك احتمال ان تكون العينتان مأخوذتان من مجتمعين لهما نفس الوسط الحسابي . أما إذا وقع الفرق بين العينتين خارج فترة الثقة الثانية نستطيع أن نؤكد ان العينتين غير مأخوذتين من مجتمعين لهما نفس الوسط الحسابي ، وبذلك يكون الفرق بينهما فرق جوهري لا يمكن أن يكون مصدره الصدفة أي اننا نستطيع أن نؤكد تأثير العامل الذي نقوم بدراسته .

مثال ٤١ :

الجدول الآتي يبين توزيع المتعطلين في منطقتين حسب أعمارهم ، والمطلوب بحث ما إذا كان هناك اختلاف جوهري بين متوسط عمر المتعطلين في المنطقتين .

العمر بالسنوات الكاملة	١١	٢٠	٣٠	٤٠	٥٠	٦٠	المجموع
عدد المتعطلين في المنطقة أ	٤	١٧	١٨	١١	٨	٥٨	
عدد المتعطلين في المنطقة ب	١٠	٢١	٢١	٢٧	١٠	٨٩	

نحسب أولاً متوسط عمر المتعطل في المنطقتين :

س-١ للمنطقة الاولى = ٣٥ر٣٤ سنة

س-٢ للمنطقة الثانية = ٣٥ر٦٧ سنة

الانحراف المعياري للعمر في المنطقة الاولى ١١ر٤٤ سنة

الانحراف المعياري للعمر في المنطقة الثانية = ١١ر٩٧ سنة

$$\therefore \text{ع ف} = \sqrt{\frac{21197}{89} + \frac{21144}{58}} = 1996 \text{ سنة .}$$

$$\therefore \text{فترة الثقة الاولى} = 1996 \pm 38416 =$$

وحيث ان الفرق بين متوسطي العينين $3534 - 3567 = 33$ سنة وهو واقع ضمن فترة الثقة الاولى .
 \therefore الفرق بينهما من المؤكد أن يكون راجعاً إلى الصدفة .

مثال ٤٢ :

من مجموع العمال المشتغلين في إحدى الصناعات اخذت عينة من العمال الذكور عددهم ٢٠٠ عامل ووجد ان متوسط اجرهم = ٣٥ ليرة بانحراف معياري ٦ ليرات ، وعينة اخرى من الاناث عددهن ٥٠ عاملة ووجد أن متوسط اجورهن ٣٠ ليرة بانحراف معياري ٣ ليرات والمطلوب اختبار الفرق بين متوسطي اجر العينين .

$$\text{ع ف} = \sqrt{\frac{9}{50} + \frac{25}{200}}$$

$$\therefore \text{فترة الثقة الاولى} = \text{صفر} \pm 1996 \times 55 = 108 \pm$$

$$\text{وفترة الثقة الثانية} = \text{صفر} \pm 2058 \times 55 = 142 \pm$$

وحيث ان الفرق بين متوسطي الأجر $35 - 30 = 5$ ليرات ، وهو يقع خارج فترة الثقة الثانية يكون في حكم المؤكد اختلاف اجور العمال الذكور والعمال الاناث اختلافاً جوهرياً .

تحليل نتائج العينات في حالة الظواهر الوصفية :

مبادئ اولية في نظرية الاحتمالات :

عندما نتكلم عن نسبة ظاهرة ما ، مثلا نسبة الامية او نسبة البطالة او نسبة المصابين بمرض معين .. الخ نعني في الواقع احتمال حدوث اي ظاهرة من هذه الظواهر . فاذا قلنا ان نسبة الامية هي ٢٥ ٪ فهل نعني بذلك ان كل ١٠٠ شخص لا بد ان يوجد بينهم ٢٥ أميين ، الواقع اننا نعني انه من كل ١٠٠ شخص يحتمل ان نجد بينهم ٢٥ اميين ، واذا عمنا هذا على المجموع يكون متوسط عدد الاميين من كل ١٠٠ شخص هو ٢٥ . وبذلك تكون النسبة ٢٥ ٪ هي متوسط لتعميم احتمال معين على مجموع معين من الوحدات . وحتى نفهم ذلك يجب ان ندرك الفرق بين المعنى الرياضي والمعنى الاحصائي للاحتمالات .

فالشخص العادي عندما يتكلم على احتمال ظاهرة ما إنما يعني عدم تأكده من حدوثها أو عدم حدوثها وهو بذلك يدلي برأيه في شكل مبهم غير واضح حيث أن رأيه هذا لا يفرق بين الاحتمال الضعيف أو القوي ، وبمعنى آخر لا يقوم هذا الرأي على أي أساس من القياس الدقيق .

على ان الشخص العادي قد يحاول اعطاء رأيه بعض الأهمية فيقول مثلا ان احتمال وقوع حدث معين هو ٩٩ ٪ إظهاراً لدرجة ثقته في وقوع الحدث او ٥ ٪ إظهاراً لدرجة عدم ثقته في وقوعه ، إلا ان ذلك لا يعني ان هذا الشخص قد استعمل مقياساً دقيقاً للتعبير عن احتمال هذا الحدث فهذا القياس ليس الا مجرد تقدير شخصي لا أساس رياضي له .

اما الاحتمال في المعنى الرياضي فيتميز عن الاحتمال بمعناه العادي السابق في انه يمكن قياسه وتحديدده بدقة إذا توفرت لدينا معلومات عن الحدث المراد تقدير احتمال وقوعه وان هذا القياس ليس قائماً على أساس شخصي وإنما هو قياس موضوعي بمعنى انه لا يمكن ان يختلف اثنان في قياس احتمال وقوع حدث معين أو عدم وقوعه إذا أقاما هذا القياس على أساس نفس المعلومات اللازم توفرها .

ويقاس الاحتمال بمقياس يبدأ بالصفر في ناحيه وبالواحد الصحيح في ناحية أخرى . ويدل الواحد الصحيح في قمة المقياس على التأكد التام من الحدث بحيث لا يكون هناك أي شك يساورنا عندما نتكلم عن هذا الحدث ، ومثلاً على ذلك احتمال أن يموت الانسان يوماً هو احتمال كامل يساوي الواحد الواحد الصحيح حيث لا مجال للشك في وقوع هذا الحدث ، وفي هذه الحالة يعبر الرياضيون عن هذا الاحتمال كالآتي : $(ح = ١)$ أما اسفل المقياس والمرقم بالصفر فيعبر عن الاستحالة التامة للحدث وفي هذه الحالة لا يكون هناك أي شك بل تأكد تام من عدم وقوع الحدث ، ومثلاً على ذلك احتمال أن يستطيع انسان ما ان يعبر سباحة المحيط الاطلسي = صفر حيث ان الفشل في المحاولة أمر مؤكد وبذلك تكون $ح = صفر$.

والواقع انه لو كانت كل الاحداث التي يحاها الانسان في حياته من الممكن البت فيها بشكل قاطع ولا مجال للشك فيها بأي حال من الاحوال لما كان هناك دافع الى دراسة موضوع الاحتمالات إلا ان الحياة تواجه الانسان بمشاكل لا حصر لها والتي لا يستطيع ان يدلي فيها برأيه بشكل قاطع بحيث تكون $ح = ١$ أو $ح = صفر$. فالطبيب مثلاً يعلم ان البنسيلين يفيد مرض معين عند مريض ما ولكنه لا يستطيع أن يؤكد ان استعماله سوف يشفي هذا المريض وبذلك لو كان الطبيب دقيقاً لما استطاع أن يقول ان احتمال شفاء المريض باستعمال البنسيلين $= ١$ أي انه يثق بذلك ١٠٠ % .

والمشكلة التي تواجهنا الآن هي كيف نستطيع ان نقيس احتمال حدث معين اذا لم يكن من النوعين السابق الاشارة اليها في الفقرات السابقة .

إذا رمينا قطعة من النقود الى أعلى وكانت كاملة التوازن وأجري رميها بدون أي تحيز فما هو احتمال أن تسقط على أحد وجهيها عندما تستقر ؟ من الواضح انه لا يمكن أن يحدث إلا أحد أمرين حيث انها اما أن تسقط على الوجه وإما أن تسقط على الظهر وبذلك يكون الاحتمالين متعادلين ، وعلى أساس جعل مقياس التأكد $= ١$ صحيح فان احتمال سقوط قطعة النقود على

أحد وجهيها $= \frac{1}{2}$ حيث اننا قسمنا الواحد بالتساوي بين الاحتمالين المتعادلين

وبمعنى آخر ان احتمال سقوط قطعة النقود على الوجه $= \frac{1}{2}$ واحتمال سقوطها

على الظهر $= \frac{1}{2}$ وبمجموع الاحتمالين $= 1$ وبديل هذا المجموع على تأكدنا التام من

سقوطها حتماً على أحد الوجهين وهذا المجموع قسمناه بالتساوي بين الاحتمالين الممكنين وقوعهما. على انه لا يجب أن نفهم من هذا اننا اذا رمينا قطعة النقود مرتين فان الوجه سرف يظهر في إحداهما وسوف يظهر الظهر في المرة الاخرى فقد نرمي القطعة عشر مرات دون أن يظهر أحد الوجهين بتاتاً . وبذلك يكون

المعنى الدقيق لقولنا ان احتمال سقوط قطعة النقود على الوجه يساوي $\frac{1}{2}$

هو اننا إذا رمينا قطعة النقود عدداً كبيراً جداً من المرات فان الوجه سوف يظهر في نصف هذه المرات (تقريباً) وسنعود الى شرح ذلك بالتفصيل فيما بعد .

وعلى نفس الأساس إذا رمينا زهرة نرد كاملة التوازن ، ودون أي تحيز

فان احتمال ظهور أي وجه لها $= \frac{1}{6}$ حيث اننا متأكدين من اظهارها لوجه

واحد من الأوجه الستة وبذلك يكون مجموع الاحتمالات التي يمكن أن تظهرها زهرة النرد $= 1$ ، وهذا المجموع يقسم بالتساوي بين الست أوجه التي تكون الزهرة (حيث اننا بالطبع ستبعد أن تقع الزهرة على أي ركن من أركانها أو اي حرف من أحرفها) .

كذلك إذا خلطنا مجموعة من ورق اللعب بحيث يكون لكل ورقة نفس الفرصة في الظهور عند السحب منه ، وإذا لم يكن الساحب متحيزاً فان

احتمال ظهور أي ورقة $= \frac{1}{52}$ حيث اننا متأكدين من ظهور ورقة واحدة من

٥٢ ورقة ولما كان مجموع الاحتمالات التي يمكن أن تظهر = ١ وهذا المجموع مقسم بالتساوي (نظرا لعدم التحيز) بين الـ ٥٢ ورقة فان احتمال ظهور ورقة واحدة = $\frac{1}{52}$. كما ان احتمال ظهور اية صورة = $\frac{12}{52} = \frac{3}{13}$ حيث انه من الـ ٥٢ احتمال هناك ١٢ احتمال كل منها يمكن أن يكون صورة .

ويمكننا أن نعبء بشكل آخر عما تقدم فنقول اننا اذا رمينا زهرة نرد ٦٠٠ مرة فانتا نتوقع ظهور وجه معين ١٠٠ مرة منها ، وتوقعنا لا يعني تأكدنا من ذلك تماما بل بالعكس فإن تعجبنا يكون شديدا لو اتفقت نتيجة التجربة مع ما نتوقعه اتفاقاً تاماً . وبذلك يمكننا أن نقيس الاحتمال بطريقة اخرى وذلك بعدد المرات التي يحدث فيها الحدث من مجموع المحاولات التي أجريناها ، على أن يكون عدد المحاولات كبيراً حتى لا يكون للصدفة أي أثر على النتيجة التي نتوصل اليها وعلى ذلك يمكن وضع القاعدة التالية :

$$\text{احتمال حدث ما} = \frac{\text{عدد مرات وقوع الحدث}}{\text{عدد المحاولات}}$$

ونضرب لذلك مثلاً : إذا قام جراح باجراء عملية جراحية معينة على ٢٠٠ شخص وتوفي ١٦ منهم نتيجة العملية نستطيع أن نقول ان احتمال الوفاة

$$\text{بسبب اجراء هذه العملية} = \frac{16}{200} = 0.08$$

ونظرية الاحتمالات مع تعقيدها تنبني على قاعدتين أساسيتين وهما جمع الاحتمالات وضربها .

جمع الاحتمالات :

إذا كان احتمال وقوع حدث معين = ح ، واحتمال وقوع حدث آخر =

ح_٢ وكان حدوث أيها مانعاً من حدوث الآخر ، فعلى أساس وقوع أي من الحدثين يكون بمثابة نجاح فان احتمال النجاح = ح_١ + ح_٢ فإذا افترضنا تقدم ١٠ أشخاص من بيروت و ٨ من الشمال و ٢ من الجنوب و ٧ من الجبل و ٩ من البقاع للحصول على وظيفة ما وقررت لجنة الاختبار سحب اسم واحد من أسماء المتقدمين يكون :

$$\frac{10}{36} = \frac{10}{9 + 7 + 2 + 8 + 10} = \text{احتمال ظهور شخص من بيروت}$$

$$\frac{8}{36} = \text{احتمال ظهور شخص من الشمال}$$

$$\frac{2}{36} = \text{احتمال ظهور شخص من الجنوب}$$

$$\frac{7}{36} = \text{احتمال ظهور شخص من الجبل}$$

$$\frac{9}{36} = \text{احتمال ظهور شخص من البقاع}$$

$$\frac{9}{36} + \frac{7}{36} + \frac{2}{36} + \frac{8}{36} = \text{أما احتمال ظهور شخص من خارج بيروت}$$

$$\frac{2}{36} + \frac{8}{36} = \text{واحتمال ظهور شخص من الجنوب والشمال وهكذا}$$

(يلاحظ ان جمع الاحتمالات يقوم على أساس ان وقوع أي حدث من الاحداث التي يمتثل وقوعها يكون مانعاً لوقوع الاحداث الاخرى ، أما إذا كان من الممكن أن يقع الحدثين في نفس الوقت فلا يجوز مطلقاً جمع احتماليهما) .

ولزيادة الايضاح نأخذ مثلاً آخر ، إذا قذفنا زهرة نرد فان احتمال ظهور أي وجه لها يكون مانعاً لظهور الواجهة الأخرى ، وبذلك يكون احتمال ظهور الواحد أو الخمسة يساوي احتمال وقوع الواحد + احتمال وقوع الخمسة = $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$ واحتمال وقوع الواحد أو الخمسة أو الثلاثة =

$$\frac{1}{2} = \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$$

وفي هذا المثال يكون حسابنا صحيح حيث يتوفر الشرط الأساسي لجمع الاحتمالات وهو ان ظهور كل حدث يكون مانعاً لظهور الاحداث الأخرى .

أما إذا رمينا قطعة من النقود إلى أعلى مرتين متتاليتين يكون احتمال ظهور الكتابة في المرة الأولى $\frac{2}{1}$ وكذلك في المرة الثانية يكون احتمال ظهور الكتابة $\frac{2}{1}$ ، فلو كان من الجائز جمع الاحتمالين لكان احتمال ظهور الكتابة في احدى المرتين $\frac{2}{1} + \frac{2}{1} = 1$ ومعنى ذلك اننا نكون متأكدين من ظهور الكتابة في احدى المرتين وهذا خطأ طبعاً لأننا نعلم جيداً انه قد نرمي قطعة النقود عدة مرات ولا نحصل على الكتابة في أي مرة منها والسبب في وقوعنا في هذا الخطأ هو تجاهل شرط التامع بين الحدثين الذي يعتبر شرطاً أساسياً في قاعدة الجمع .

ضرب الاحتمالات :

ان احتمال النجاح المزدوج لحدثين مستقلين عن بعضها يساوي حاصل ضرب احتمال الحدث الاول في احتمال الحدث الثاني ، وبديهي اننا في هذه الحالة لا نشترط أن يكون احد الحدثين مانعاً للآخر كما فعلنا في حالة جمع

الاحتمالات حيث اننا نبحث هنا في احتمال وقوع الحدثين معاً . كذلك إذا كان هناك ثلاث احداث مستقلة عن بعضها واحتمالاتها $ح١$ ، $ح٢$ ، $ح٣$ على الترتيب فان احتمال حصولها جميعاً في وقت واحد $= ح١ \times ح٢ \times ح٣$. وهكذا مهما تعددت الاحداث يكون احتمال حصولها جميعاً يساوي حاصل ضرب احتمالاتها في بعضها . ويمكننا أن نلاحظ انه لما كان احتمال كل حدث أقل من الواحد الصحيح فان ضرب الاحتمالات يؤدي إلى جعل الاحتمال الناتج أصغر ، وهذا بالطبع أمر معقول حيث ان احتمال حصول حدثين معاً أقل من احتمال وقوع حدث واحد فقط .

مثال :

إذا رمينا زهرتين نرد فما هو احتمال أن نحصل على ستة من كل منهما سوياً .

$$\frac{1}{6} = \text{احتمال النجاح في الزهرة الاولى}$$

$$\frac{1}{6} = \text{احتمال النجاح في الزهرة الثانية}$$

$$\frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = \text{احتمال النجاح المزدوج}$$

المعنى الاحصائي للاحتمال :

من المناقشة السابقة يتبين لنا كيف يمكن قياس الاحتمال رياضياً فنقول

مثلاً ان احتمال ظهور الستة عند رمي زهرة واحدة $= \frac{1}{6}$ ، غير انه بالتأمل

في هذا المقياس نجد انه يوقعنا في حيرة شديدة ذلك ان الحدث اما ان يقع أو لا يقع ولا يمكن أن يقع بخمسه أو سدسه . فعند رمي زهرة النرد فاما أن يظهر الوجه ستة او لا يظهر وبذلك بعد انتهاء التجربة نجد أن الواقع يخالف الاحتمال سواء كنا قد قدرنا هذا الاحتمال ب $\frac{1}{4}$ أو $\frac{1}{6}$ أو أي رقم آخر

قبل اجراء التجربة فالنتيجة واحدة بعد اجراء التجربة وهي اما ان يقع الحدث أو لا يقع بغض النظر عن التقدير الذي كنا قد أعطينا للاحتمال من قبل . أما المعنى الاحصائي للاحتمال فانه يعطينا فكرة اكثر وضوحاً من المقياس الرياضي . فالاحتمال بالمعنى الاحصائي هو تكرار نسبي فاذا كان

الاحتمال الرياضي $= \frac{1}{4}$ فمعنى ذلك احصائياً اننا اذا كررنا التجربة عدد كبير من المرات (بقرب من ∞) فان الحدث يقع في نصف هذه المرات ولا يقع في النصف الآخر . فاذا رمينا قطعة نقود ١٠٠٠٠٠ مرة فان النتيجة لا يمكن أن تختلف كثيراً عن ٥٠٠٠٠ وجه و ٥٠٠٠٠ ظهر الا اذا كانت القطعة ليست كاملة التوازن .

على ضوء هذا المعنى نعود للتوزيع التكراري لعدد مرات النجاح الذي سبق الإشارة اليه ولايجاد متوسط عدد مرات النجاح فنضرب كل قيمة (من قيم عدد مرات النجاح) في الاحتمال الخاص بها ونقسم على مجموع التكرارات (مجموع الاحتمالات) وهو يساوي ١ فينتج ان متوسط قيم مرات النجاح يساوي $n \times$ ح أي عدد المحاولات مضروباً في احتمال الكسب في أية محاولة باعتبار ان نجاح أي محاولة لا يؤثر في احتمال نجاح أو فشل أي محاولة أخرى وبفرض ان احتمال النجاح واحد في كل المحاولات .

مثال :

رمى لاعب ١٢ زهرة دفعة واحدة وكرر هذه التجربة عدة مرات فما هو متوسط عدد ما يحصل عليه من الشيشات لكل مرة .

$$\frac{1}{4} = \text{احتمال رمي الشيش من أي زهرة}$$

ورمي ١٢ زهرة دفعة واحدة يعني عمل ١٢ محاولة .

$$\text{متوسط عدد ما يحصل عليه من الشيشات في كل رمية} = ١٢ \times \frac{1}{4} = ٣$$

والمقصود بقولنا ان الوسط الحسابي لعدد الزهرات الناجحة = ٣ هو اننا اذا قمنا فعلاً بتجربة قد نحصل على اكثر أو أقل من زهرتين وربما نحصل على اثنتين بالضبط ولو كررنا التجربة عدداً كبيراً من المرات فان المتوسط يكون ٣ أو قريباً منها .

وكما أوجدنا الوسط الحسابي للتوزيع التكراري لعدد مرات النجاح يمكننا أن نوجد الانحراف المعياري لهذا التوزيع وهو يقيس لنا درجة الاختلاف بين عدد مرات النجاح التي يمكن ان نحصل عليها والوسط الحسابي . ويمكن ايجاد الانحراف بضرب مربع كل قيمة من قيم عدد مرات النجاح في الاحتمال الخاص بها ثم نطرح مربع الوسط الحسابي من مجموع حواصل الضرب فينتج مربع الانحراف المعياري (التباين) ويمكن اثبات ان الانحراف المعياري للتوزيع التكراري لعدد مرات النجاح
$$= \sqrt{n \cdot h \cdot l}$$

مثال :

إذا اعتبر ظهور جهار أو أكثر من زهرة الزرد نجاحاً فما هو متوسط عدد الزهرات الناجحة إذا رمينا ١٢ زهرة وما هو الانحراف المعياري من هذا المتوسط .

$$\text{احتمال النجاح في هذه الحالة بالنسبة لكل زهرة} = \frac{3}{4} = \frac{1}{2}$$

متوسط عدد ما نحصل عليه من زهرات ناجحة في كل مرة =

$$12 \times \frac{1}{2} = 6 \text{ زهرة}$$

$$\frac{1}{2} = \frac{1}{2} - 1 = \text{واحد احتمال الفشل} \quad \frac{1}{2} = \text{احتمال النجاح}$$

$$\sqrt{12 \times \frac{1}{2} \times \frac{1}{2}} = \sqrt{3} = 1.732 \text{ تقريباً ع.٠}$$

مثال :

رمى قطعة نقود كاملة التوازن ٩٠٠ مرة فالنتيجة التي نتوقعها نظرياً هي أن نحصل على وجه القطعة نصف عدد المرات تقريباً ونصل إلى هذه النتيجة بإيجاد الوسط الحسابي لعدد مرات النجاح (يعتبر ظهور وجه قطعة النقود نجاحاً) وهو في هذا المثال $900 \times \frac{1}{2} = 450$ نجاحاً في كل مرة نعيد فيها التجربة فعوامل المصادفة يمكن أن تجعل النتائج تختلف قليلاً من ذلك فقد نحصل على ٤٥٥ نجاحاً أو ٤٤٨ أو ٤٤٦ أو ٤٥٢ وهكذا . الا اننا نستطيع أن نقول اننا إذا أعدنا التجربة عدداً كبيراً جداً من المرات فان المتوسط لعدد مرات النجاح في كل التجارب التي نجربها سوف يكون ٤٥٠ .

الا أننا لا نستطيع أن نؤكد أن الفرق بين عدد مرات النجاح الواقعي والعدد المتوقع (الوسط الحسابي) يكون دائماً راجعاً الى عوامل الصدفة ، فقد تكون قطعة النقود ليست كاملة التوازن وبذلك يكون هناك تحيزاً اما نحو وجه القطعة أو ظهرها ، كذلك يمكن أن يتسرب عنصر التحيز الى طريقة رمي قطعة النقود . وبالطبع لو كان ليس بإمكاننا أن نميز بين الأخطاء (الفرق بين الرقم الواقعي والرقم المتوقع) التي يمكن أن تكون راجعة الى الصدفة وتلك التي ترجع الى عوامل أخرى كالتحيز أو أي عامل آخر يؤثر في

نتيجة التجربة فان ما نحصل عليه من نتائج يكون معدوم القيمة حيث يكون دائماً موضع الشك والتساؤل، هل تختلف النتيجة التي نحصل عليها عن النتيجة المتوقعة اختلافاً يمكن أن نعزوه الى الصدفة أو هو اختلاف جوهري راجع الى عوامل اخرى ، يمكن أن يكون التحيز احداها .

على انه احصائياً يمكن التمييز بين الفروق التي تكون راجعة الى عوامل الصدفة وتلك التي تكون راجعة الى عوامل اخرى وذلك باستخدام خواص التوزيع التكراري لعدد مرات النجاح . إذ لو كررنا المحاولة عدداً كبيراً جداً فان نتائج هذه المحاولات تتوزع توزيعاً معتدلاً تقريباً متوسطة = \bar{N} ح وانحرافه المعياري \sqrt{N} ح ل . وبذلك يلعب الانحراف المعياري للتوزيع دوراً هاماً في تحديد ما اذا كان الفرق راجعاً الى الصدفة أو غيرها . فاذا كان الفرق بين النتيجة التي نحصل عليها فعلاً والنتيجة المتوقعة أقل من الانحراف المعياري \sqrt{N} ح ل فمن المحتمل أن يكون هذا الفرق راجعاً إلى عوامل الصدفة ، أما اذا كان الفرق يعادل أو يزيد عن ضعف الانحراف المعياري فان احتمال حدوث الفرق بطريقة المصادفة يكون ضئيلاً ، وكذلك اذا كان الفرق يعادل أو يزيد عن ثلاث أمثال الانحراف المعياري يكاد يكون من المستحيل عملياً أن يكون راجعاً الى عوامل الصدفة . على أن هناك خلاف بين الاحصائيين فيما يتعلق بالتمييز بين الفروق الظاهرية (الراجعة الى الصدفة) والفروق الجوهرية (التي ترجع الى عوامل غير الصدفة) ، بعضهم يعتبر الفرق الذي يعادل $\sqrt{2}$ \sqrt{N} ح ل فرقاً جوهرياً لا يمكن أن يعزى الى عوامل الصدفة والبعض الآخر يشترط لاعتبار الفرق جوهرياً أن يكون مساوياً $\sqrt{3}$ \sqrt{N} ح ل على الأقل .

مثال ٤٣ :

انتجت احدى مزارع البرتقال ١٠٠٠٠ ثمرة وجد بينها ٥٧٠ ثمرة عطبة

بينما كانت النسبة الممنوعة للعطب هي ١٠٪ هل يدل الفرق بين عدد الثمار العطبة الذي نتوقعه نظرياً وذلك الذي حصلنا عليه فعلاً على حدوث تغير جوهري في ظروف الزراعة أم يمكن أن نغزوه الى عوامل المصادفة .

$$\text{نسبة الثمار العطبة} = \frac{750}{10000} \times 100 = 7.5\%$$

والمطلوب معرفة هل الفرق بين ٧.٥٪ وهي النسبة الواقعية و ١٠٪ وهي للنسبة المتوقعة فرق ظاهري أو فرق جوهري .

عدد الثمار التي يتوقع عابها $= \frac{1}{10} \times 10000 = 1000$ ثمرة (الوسط الحسابي) .

$$\text{الانحراف المعياري} = \sqrt{\frac{9 \times 1 \times 10000}{10 \times 10}} = 30$$

الفرق بين عدد الثمار العطبة الواقعي والمتوقع $= 1000 - 570 = 430$ ثمرة.

$$\text{هذا الفرق} = \frac{430}{30} = 14.33 \text{ ع}$$

وهو فرق يزيد كثيراً على ثلاث أمثال الانحراف المعياري فهو اذن فرق جوهري يدل على تحسن فعلي في ظروف الزراعة .

الوسط الحسابي والانحراف المعياري لنسبة النجاح :

في الفقرات السابقة ناقشنا الوسط الحسابي والانحراف المعياري لعدد مرات النجاح ، غير أنه من الأفضل أحياناً أن نعبّر عن مرات النجاح في شكل

نسبة مئوية بدلاً من التعبير عنها في شكل اعداد مطلقة . وفي هذه الحالة يكون الوسط الحسابي لنسبة النجاح $\frac{ن ح}{ن} = ح$ والانحراف المعياري

للنسبة = $\sqrt{\frac{ن ح ل}{ن^2}}$ أي $\sqrt{\frac{ح \times ل}{ن}}$ وعند استعمال هذه القوانين يجب أن

تلتبه إلى أن ح و ل فيها ليست نسب مئوية بل نسب منسوبة الى الواحد الصحيح . وعلى أساس هذه القوانين يمكن حل امثال السابق كالاتي :

متوسط نسبة النجاح $\frac{1}{1.0} = 10\%$ أي

$$ع = \sqrt{\frac{\frac{9}{10} \times \frac{1}{10}}{10000}} = 0.003$$

ع للنسبة المئوية $ع = 100 \times 0.003 = 3$

الفرق بين النسبة المتوقعة والنسبة الفعلية بدلالة ع $1433 = \frac{570.5}{3}$

وهو فرق أكبر بكثير من 3 أمثال 3 فهو اذن فرق جوهري .

ويمكن أن نصل إلى نفس النتيجة إذا أجرينا حسابنا على النسبة الواقعية

0.57 %

$$ع = \sqrt{\frac{\frac{430}{1000} \times \frac{570}{1000}}{10000}} = 0.0049$$

ع للنسبة المئوية $ع = 100 \times 0.0049 = 49$

$$\text{الفرق بين النسبة الواقعية والمتوقعة بدلالة ع} = \frac{58010}{7049} = 77$$

وهو فرق أكبر بكثير من ٣ أمثال ٠٤٩ ر. فهو اذن فرق جوهري .

مثال ٤٤ :

انبت احد أنواع البسلة ١٤٣٠٠ من البذور الصفراء و ٤٩٠٠ من البذور الخضراء في حين ان عدد البذور الصفراء الذي نتوقعه حسب قانون مندل يعادل ٧٥ ٪ من مجموع البذور ، هل يمكن أن نعزو الفرق بين العدد الذي حصلنا عليه فعلا وذلك الذي نتوقعه نظرياً إلى عوامل المصادفة :

$$\text{عدد البذور الصفراء المتوقعة} = \frac{3}{4} \times 19200 = 14400$$

$$\text{الفرق بين العدد المتوقع والعدد الفعلي} = 14300 - 14400 = 100$$

$$\text{ع} = \sqrt{\text{ن ح ل}} = \sqrt{\frac{1}{4} \times \frac{3}{4} \times 19200} = \sqrt{3600} = 60$$

$$\text{الفرق بدلالة ع} = \frac{100}{60} = 1.66$$

وهو أقل من ٣ أمثال ع فهو اذن فرق ظاهري يمكن أن نعزوه إلى المصادفة أي ان النتيجة لا تعتبر مخالفة لقانون مندل .

ويمكن اجراء العمل بالطريقة الآتية :

$$\text{ع} = \sqrt{\frac{\frac{1}{4} \times \frac{3}{4}}{19200}} = 0.031$$

$$\text{ع للنسبة المئوية} = 100 \times 0.031 = 3.1$$

$$\text{الفرق بين النسبتين} = ٧٥ - ١٠٠ \times \frac{١٤٣٠٠}{١٩٢٠٠} = ٥$$

هذا الفرق أقل بكثير من ٣ امثال ع (٩٢ ر) ، وهو بذلك فرق يمكن أن يكون راجعاً إلى الصدفة

ويمكن اجراء العمل بالطريقة الآتية :

$$ع = \sqrt{\frac{\frac{٤٩}{١٩٢} \times \frac{١٤٣}{١٩٢}}{١٩٢٠٠}} = ٠٠٣١ ر$$

$$ع \text{ للنسبة المئوية} = ٠٠٣١ ر \times ١٠٠ = ٣١$$

$$\text{الفرق بين النسبتين} = ٧٥ - ٧٤ ر٥ = ٥ ر بدلالة ع = \frac{٥}{٣} = ١ ر٦٦$$

أي أقل من ٣ ع فهو اذن فرق يمكن أن يكون راجعاً إلى الصدفة .

تطبيق هذه النتائج في دراسة العينات :

عندما نسحب عينة من مجتمع ما بواسطة الارقام العشوائية فانها تكون احدى عينات كثيرة (او محاولات كثيرة) يمكن أن تسحب من هذا المجتمع وإذا كان موضوع دراسة العينة هو احتمال ظاهرة ما (احتمال البطالة أو احتمال الوفاة او احتمال النجاح الخ ..) فاننا نستطيع أن نكون توزيعاً تكرارياً لعدد العينات التي تعطي كل احتمال من احتمالات الظاهرة موضوع الدرس وقد رأينا ان التوزيع التكراري لهذه الاحتمالات المختلفة عندما يكون عدد المحاولات كبيراً جداً (عدد العينات) يتخذ شكل

منحنى طبيعي ويكون الانحراف المعياري لهذا التوزيع بدلالة الاحتمال مساوياً $\sqrt{n \cdot h \cdot l}$ وبشكل آخر $\sqrt{\frac{h \cdot l}{n}}$ إذا أردنا قياس

الانحراف المعياري لاحتمال النسبة . وعلى هذا الأساس نستطيع أن نحسب مدى الانحراف بين الاحتمال المقدر من العينة والاحتمال العام للمجتمع الاصلي الذي يمثله متوسط توزيع الاحتمالات الخاصة بالعينات المختلفة المسحوبة من هذا المجتمع .

وفي اجراء الاختبارات المختلفة بالنسبة للظواهر الوصفية على أساس العينات سوف نتبع نفس الخطوات التي سبق أن أثمنا اليها بالنسبة للظواهر الرقمية (المتوسطات) وبذلك لا يكون هناك داعي لشرح خطوات العمل ونكتفي لذلك باعطاء امثلة مع نماذج حلولها .

تقدير نسبة ظاهرة ما في مجتمع معين من واقع بيانات العينة :

مثال ٤٥ : اجريت دراسة بالعينة تبين منها ان ٧٠٠ شخص مرضى بمرض معين من مجموع عدد الأشخاص في العينة والبالغ عددهم ٢٥٠٠ شخص . المطلوب تقدير نسبة الاصابة بهذا المرض في المجتمع .

$$\text{نسبة الاصابة في العينة} = \frac{700}{2500} \times 100 = 28\%$$

$$\frac{\sqrt{0.28 \times 0.72}}{2500} \sqrt{\frac{h \cdot l}{n}} = \text{الخطأ المعياري لنسبة الاصابة (للوحدة الواحدة)}$$

$$= 0.0089$$

الخطأ المعياري للنسبة المئوية للاصابة = $0.0089 \times 100 = 0.89$

∴ فترة الثقة بدرجة ثقة ٩٥ % = $0.89 \times 1.96 = 1.74$

∴ نسبة الاصابة في المجتمع بدرجة ثقة ٩٥ % تتراوح بين ٢٨ % ±

$$1.74\% = 29.74\% , 26.26\%$$

اختبار معنوية الفرق بين نسبة ظاهرة ما في مجتمع معين وفي العينة :

مثال ٤٦ : في عام ١٩٤٢ كان عدد المواليد في الجمهورية العربية المتحدة ٦٥٨٣٢٤ منهم ٣٤٣٦٣٣ ذكر وفي نفس العام كان عدد المواليد في احدى المحافظات ٣٢٢٧٧ منهم ١٦٦٨٣ ذكر - فهل يستدل من ذلك على وجود اختلاف جوهري في نسبة الذكور من المواليد ؟

$$\frac{343633}{658324} = \text{احتمال أن يكون المولد ذكراً في القطر كله}$$

$$= 0.522 \text{ تقريباً}$$

$$\frac{16683}{32277} = \text{احتمال أن يكون المولد ذكراً في المحافظة}$$

$$= 0.516 \text{ تقريباً}$$

الخطأ المعياري لاحتمال ان يكون المولود ذكراً =

$$\sqrt{\frac{0.516 \times 0.484}{32277}} = 0.027$$

$$\text{الفرق بدلالة } E = \frac{0.516 - 0.522}{0.027} = 2.22 \text{ تقريباً}$$

وبذلك يكون الفرق واقعاً ضمن درجة ثقة ٩٩ %

∴ يمكن أن يكون هذا الفرق راجعاً الى الصدفة .

ويمكن الوصول إلى نفس النتيجة بطريقة أخرى حيث يكون عدد المواليد الذكور المتوقع في الجيزة على اساس الاحتمال العام $= 0.522 \times 32277 = 17849$ تقريباً .

وبذلك يكون الفرق بين العدد المتوقع والعد الحقيقي =

$$166 = 16683 - 17849$$

الخطأ المعياري لعدد المواليد المتوقع =

$$\sqrt{0.522 \times 0.478 \times 32277} = 89 \text{ تقريباً}$$

$$\frac{166}{89} = 1.86 = \text{الفرق بدلالة ع}$$

وبذلك يكون الفرق راجعاً الى الصدفة .

اختبار معنوية الفرق بين عينتين :

مثال ٤٧ :

في عينة من ٣٠٠ أسرة كانت نسبة الاسر التي يزيد دخلها عن ١٥ دينار هي ٣٥٪ وفي عينة أخرى من ٢٥٠ أسرة كانت النسبة ٣٠٪ -
اختبر معنوية الفرق بين العينين : -

$$\text{الفرق الواقعي بين النسبتين} = 35 - 30 = 5\%$$

$$\frac{\frac{0.07 \times 0.3}{250} + \frac{0.065 \times 0.35}{300}}{\sqrt{}} = \text{ع ف}$$

$$= 0.04$$

$$\text{ع ف للنسبة المئوية} = 0.04 \times 100 = 4$$

$$\text{بدرجة ثقة } 95\% \text{ يكون مدى الثقة صفر} + 5 \times 1.96 + 7.84$$

وحيث أن الفرق 5% أقل من هذا الفرق المحتمل ، إذا الفرق بين العينتين يمكن أن يكون راجعاً الى الصدفة .

حل آخر

نفترض أن العينين تنتميان لمجتمعين متشابهين من حيث نسبة الأمر التي يزيد دخلها عن ١٥ دينار ، أي لهما نفس نسبة هذه الأمر ، وحيث أن هذه النسبة العامة لهذين المجتمعين غير معروفة فاننا نقدرها كوسط حسابي مرجح للنسبتين في العينين .

$$\text{أي} = \frac{250 \times 30 + 300 \times 0.35}{550} = 0.33$$

ويكون الخطأ المعياري لمجتمع الفروق بين العينات العشوائية التي يمكن

$$\sqrt{\left(\frac{1}{250} + \frac{1}{300} \right) 33 \times 33} = \text{ع ف}$$

$$= 0.04$$

وبذلك يكون الخطأ المعياري للنسبة المئوية بدرجة ثقة 95% .

$$= 0.04 \times 100 \times 1.96 + 8.74$$

وحيث ان الفرق بين العينتين اقل من هذا الفرق المحتمل ، اذا الفرق بين العينتين يمكن ان يكون راجعاً الى الصدفة .

لوحات مراقبة الانتاج :

اشرنا سابقاً الى كيفية اجراء المراقبة الاحصائية للانتاج على اساس تصميم لوحات تبين عليها فترات الثقة المطلوبة ثم يجري الاختبار على وحدة واحدة من الانتاج ويعين موقعها في لوحة المراقبة ، ثم بعد فترة من الزمن تؤخذ وحدة اخرى ونبين موقعها في لوحة المراقبة وهكذا يجري العمل على اساس وحدات كل وحدة منها قائمة بذاتها . وحيث انه في الغالب تجري المراقبة ليس على وحدات ، كل وحدة على حدة ، وانما على اساس عينات من الانتاج فيجدر بنا ان ندرس كيف يتم العمل في هذه الحالة .

وفكرة المراقبة في هذه الحالة لا تختلف بتاتا عن الفكرة التي سبق الاشارة اليها ، حيث ان جميع العينات التي يحتمل أخذها من الانتاج تكون توزيعاً معتدلاً يكون متوسطه هو المتوسط العام للظاهرة موضوع المراقبة ونسبته هي النسبة العامة للظاهرة موضوع المراقبة . واذا كانت الفروق بين العينات والمجتمع الذي سحبت منه ناتجة عن عوامل الصدفة تكون مقاييسها واقعة ضمن فترات الثقة التي حددناها على لوحة المراقبة ، أما إذا كانت خارجة عن هذه الفترات فيكون الفرق بينها وبين المجتمع الذي سحبت منه فرقاً جوهرياً ناتجاً عن أي عامل آخر غير الصدفة ، وقد يكون خلافاً في آلات الانتاج أو اهمالاً من العامل أو عدم انتظام الآلات في عملها . ومن الواضح أن الانحراف المعياري لتوزيع العينات (الخطأ المعياري) يحسب بالقانون

$$\frac{\sigma}{\sqrt{n}} \quad \text{إذا كنا بصدد درس متوسط ظاهرة معينة أو بالقانون}$$

إذا كنا بصدد درس نسبة ظاهرة معينة .

مثال ٤٨ :

يستخدم مصنع ما آلة تنتج انابيب متوسط قطرها ٠.٥٧٤ بوصة بانحراف معياري ٠.٠٠٨ بوصة . صمم لوحة لمراقبة انتاج هذه الآلة على أساس عينات تتكون كل منها من ٦ انابيب كل ساعتين عمل .

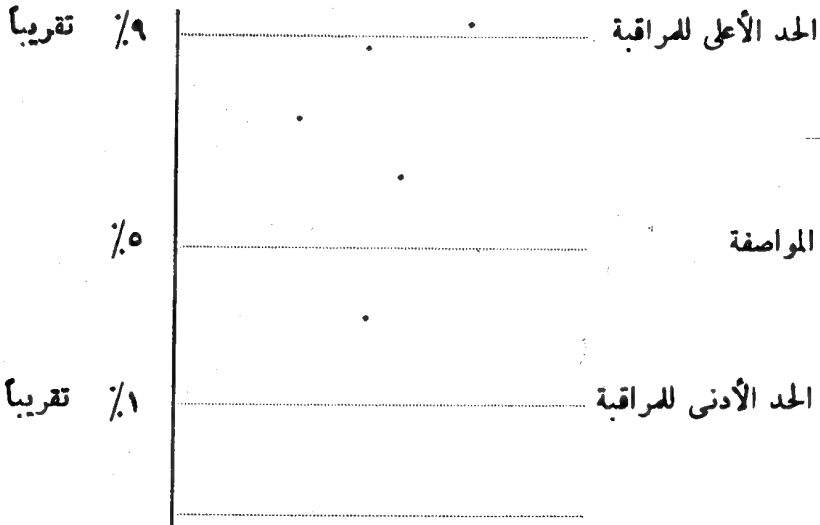
$$\text{الخطأ المعياري} = \frac{\sqrt{٠.٥ \times ٠.٥}}{٢.٠٠} = ٠.١٥٤$$

بدرجة ثقة ٩٩٪ يكون حدى المراقبة بين

$$\underline{+٠.٥} + ١.٠٥٤ \times ٢.٥٨$$

$$= ٣.٩٧٣٢ + ٠.٥$$

أي بين ٨.٩٧٣٢ ، ١.٠٦٨



واضح من الرسم ان النقطة الاخيرة تدل على ان نسبة التلف قد زادت الى

درجة لا يمكن ان يكون الفرق عندها راجعة الى الصدفة ، فلا بد لذلك من الكشف على اسباب هذا الفرق الجوهرى .

تحديد حجم العينة :

ان تحديد حجم العينة في بحث ما امر هام جداً حيث يتوقف عليه تحديد التكاليف التي يتطلبها البحث والمجهودات التي سوف تبذل فيه . ولا يمكن اجراء هذا التحديد الا بالاجابة على الاسئلة الآتية :

١ - ما هو الخطأ الذي نكون مستعدين لقبوله في البحث ؟ ولا شك ان ذلك يتوقف بالدرجة الاولى على الغرض الذي يجري من اجله البحث أي على الناحية التي سوف تستخدم فيها النتائج . ومن الواضح ان بعض البحوث تتطلب دقة كبيرة الا ان ذلك يعني ان حجم العينة لا بد ان يكون كبيراً ، الأمر الذي يؤدي الى رفع التكاليف ، وبعض البحوث الاخرى لا تتطلب ذلك وبذلك يمكن قبول خطأ أكبر لانقاص التكاليف والمجهود .

٢ - ما هي درجة الثقة التي نرغب في تحقيقها من ناحية الخطأ ؟ ومن الواضح ان زيادة درجة الثقة تتطلب حجماً كبيراً للعينة ؛ والعكس صحيح ، وبذلك يرتبط الامر ايضاً بتكاليف البحث والمجهود التي يمكن توفيرها له .

٣ - لا بد من معرفة معالم المجتمع الذي نريد أن نأخذ عينة منه ، وتتوقف هذه المعالم على نوع الدراسة المطلوب إجرائها ، فاذا كنا بصدد بحث عن متوسط ظاهرة ما فاننا نحتاج بذلك أن نتعرف على تباين المجتمع بالنسبة لهذه الظاهرة ، وإذا كنا بصدد بحث عن نسبة ظاهرة ما لا بد أن يكون لدينا معرفة تقريبية من هذه النسبة .

٤ - إذا كنا ندرس عينة طبقية فانه يمكن حساب حجم العينة لكل قسم من المجتمع وبالجمع يمكن حساب حجم العينة الكلي .

هـ - يمكن أن يكون البحث بالعينة لدراسة أكثر من ظاهرة واحدة في المجتمع ، وحجم العينة الذي يصلح لدراسة ظاهرة منها قد لا يكون صالحاً لدراسة باقي الظواهر . وفي هذه الحالة نحدد الأخطاء حسب الظواهر الأكثر أهمية في الدراسة ونحدد حجم العينة الذي يناسب كل من هذه الظواهر المنفصلة فإذا كانت الاحجام متقاربة فنأخذ أكبر حجم منها طالما كان في حدود النفقات والجهود التي نستطيع أن نتحملها . أما إذا كانت الاحجام متفاوتة كثيراً فلا بد أن نضحي ببعض الدقة بالنسبة لبعض الظواهر حتى نجعلها متقاربة وتأخذ الحجم الأكبر . أما إذا كان التفاوت من الكبر بحيث لا يمكن مع تضحيه بعض الدقة ان نجعلها متقاربة ففي هذه الحالة لا بد أن نهمل بعض الظواهر من البحث .

فإذا استطعنا أن نعرف معالم المجتمع وأن نحدد درجة الخطأ التي نقبلها ودرجة الثقة في هذا الخطأ فان تحديد حجم العينة يكون مجرد استنتاج جبري من إحدى معادلات الخطأ المعياري (تبعاً لنوع الدراسة) . والخطأ المقبول قد يكون في صيغة خطأ مطلق أو في صيغة معامل اختلاف لهذا الخطأ أي في صيغة نسبة مئوية من التقدير المطلوب .

مثال ٥٠ :

في مجتمع حجمه ٣٠٠ وحدة وجد ان متوسطه = ١٩ وتباينه = ٨٥٦
أوجد عدد الوحدات التي يجب سحبها في عينة عشوائية لتقدير المتوسط بحيث يكون الخطأ ١٠ ٪ من التقدير باحتمال ٩٥ ٪ .

حيث ان الخطأ معطى في شكل معامل اختلاف وحيث ان درجة الثقة المطلوبة هي ٩٥ ٪ يكون قانون الخطأ المعياري الذي نستعمله هو :

$$\frac{\frac{ع}{2}}{\sqrt{\frac{2}{3}}} = ع \quad (استخدمنا 2 تقريباً) \quad 100 \times \frac{100}{-س}$$

$$\frac{\frac{856}{2} \sqrt{2}}{\sqrt{\frac{2}{3}}} = 10 \therefore 100 \times \frac{10}{19}$$

$$100 \times 100 \times \frac{1}{19} \times \frac{1}{19} \times \frac{1}{3} \times \frac{856}{10} \times 4 = 100 \text{ بالتربيع}$$

$$\frac{342400}{23610} = 1$$

$\therefore ن = 95$ تقريباً .

(نلاحظ انه حيث ان حجم المجتمع محدد ، كان من الأفضل من ناحية

$$\frac{\frac{ع}{2} \sqrt{\frac{2}{3}}}{\frac{2-ن}{1-ن}} = ع \text{ لدقة استخدام القانون}$$

(حيث ن حجم المجتمع ، 2 حجم العينة) إلا أن الفرق في النتيجة لا يكون من الأهمية بمكان كبير) .

مثال ٥١ :

إذا كان متوسط الدخل في المجتمع ٢٥ ديناراً ، وتباينة ٢٢٥ ديناراً .

كم يكون حجم العينة اللازمة للحصول على معاميل اختلاف للمتوسط ١٠٪ .

$$100 \times \frac{\frac{e}{\sqrt{2}}}{s} = e$$

$$100 \times \frac{\frac{e}{\sqrt{2}}}{25} = 100$$

$$10000 \times \frac{1}{625} \times \frac{225}{2} = 100 \text{ بالتربيع}$$

$$22500 = 2 \times 625 \therefore$$

$$\therefore 36 = 2 \text{ وحدة .}$$

مثال ٥٢ :

كم يكون حجم العينة لدراسة نسبة انتشار مرض ما اذا كان من المعروف ان هذا المرض ينتشر في العادة بنسبة ١٪ تقريباً وإذا كان الخطأ المطلوب هو ٢٠٪ من التقدير . كم يكون حجم العينة اذا كان الخطأ المطلوب هو بالنسبة لكل من الذكور والاناث .

حيث ان الدراسة خاصة بنسبة ظاهرة ما نستخدم قانون الخطأ المعياري للنسبة في شكل معامل اختلاف حيث ان الخطأ معطى في شكل نسبة مئوية من التقدير .

$$\sqrt{\frac{c(c-1)}{d}}$$

$$c = \frac{100 \times (حيث ل = c - 1)}{c}$$

$$\sqrt{\frac{99 \times 101}{d}}$$

$$20\% = 100 \times \frac{1}{5}$$

$$\text{بالتربيع } 400 = \frac{1}{100} \times \frac{99}{100} \times \frac{1}{d} \times \frac{100}{1} \times \frac{100}{1} \times 10000$$

$$\frac{9900}{d} = 400$$

$$24750 \div$$

$$9900 = 24 \div$$

ونفس الحجم نحتاجه لكل من الذكور والاناث أي أن العينة في مجموعها يجب أن تتكون من ٩٥٠ نصفهم من الذكور والنصف الآخر من الاناث .

وواضح اننا اذا أردنا خطأ أقل بنسبة ١٠٪ فقط مثلاً فان حجم العينة لكل من الذكور والاناث يجب أن يكون ٩٩٠٠ وحدة أي أن العينة في مجموعها يجب أن تكون ١٩٨٠٠ وحدة . أما اذا قبلنا خطأ أكبر فان حجم العينة يقل . واذا كنا نريد نفس الخطأ السابق ٢٠٪ ولكن بدرجة ثقة ٩٥٪

فان حجم العينة يجب أن يزيد . ولهذا بعد تحديد حجم العينة يجب أن نعلم ما اذا كان هذا الحجم يتناسب مع التكاليف والجهود التي في استطاعتنا تخصيصها للدراسة أم لا ، فإذا كان يتناسب معها أجرينا البحث بالحجم المحدد ، أما اذا كان يفوق هذه التكاليف والجهود علينا أن نقنع بدرجة دقة أقل أي بخطأ معياري أكبر ، أو نعمل على اجراء آخر نستطيع به أن ننقص من الانحراف المعياري للمجتمع وذلك باتباع نوع آخر من العينات حيث قد يكون من الممكن تقسيم هذا المجتمع الى طبقات متجانسة وبذلك ينقص الانحراف المعياري فستطيع بذلك انقاص حجم العينة مع الابقاء على الدقة المطلوبة .

ومن ناحية الدراسات الخاصة بالنسب نلاحظ أن حجم العينة يتناسب تناسباً عكسياً مع احتمال وجود الظاهرة في المجتمع ، اذا كلما كان الاحتمال صغيراً أي كلما كانت النسبة صغيرة كلما كان حجم العينة التي يجب أخذها كبيراً (الشيء النادر يحتاج الى بحث أوسع وأعمق) ، والعكس كلما كانت الظاهرة شائعة في المجتمع كلما كان في استطاعتنا أن نأخذ عينة صغيرة ونحقق نفس درجة الدقة المطلوبة ، ففي المثال السابق اذا كان المرض منتشراً بنسبة ١٠٪ وليس ١٪ فقط فان حجم العينة التي يجب أخذها يكون :

$$100 \times \frac{\frac{\sqrt{1 \times 9}}{2}}{1} = 20$$

$$10000 \times \frac{10}{1} \times \frac{10}{1} \times \frac{1}{2} \times \frac{9}{100} \times \frac{1}{100} = 400 \therefore$$

$$900 = 24 \therefore$$

$$225 = 2 \therefore \text{وحدة}$$

تمارين

- ١ - اذا علمت ان متوسط عمر اللعبة الكهربائية حسب عينة مكونة من ١٠٠ لمبة هو ١١٩٠ بانحراف معياري ٩٠ ساعة ، ومن عينة اخرى مكونة من ٧٥ لمبة ظهر ان المتوسط هو ١٢٣٠ ساعة بانحراف معياري ١٢٠ ساعة . اختبر معنوية الفرق بين العينين .
- ٢ - ينتج احد المصانع جوارب نسبة التلف فيها ٣٪ - صمم لوحة لمراقبة الانتاج على أساس عينات تتكون كل منها من ٢٠٠ جورب .
- ٣ - ينتج احد المصانع خيطان قوة تحملها ٨٧٦٤ أوقية بانحراف معياري ١٢٨ أوقية - صمم لوحة لمراقبة الانتاج على اساس عينات تتكون كل منها من ١٦ خيط .
- ٤ - تبين من عينة عشوائية مكونة من ١٠٠ عامل ان متوسط الأجر هو ٢٧٥ ليرة بانحراف معياري ٥ ليرات - قدر متوسط الأجر لعمال الصناعة كلها .

الفصل الحادي عشر

الارتباط

يقصد بالارتباط بين ظاهرتين أو بين متغيرين وجود علاقة بينهما بمعنى انه اذا تغير أحد المتغيرين في اتجاه معين (بالزيادة أو بالنقص) فان المتغير الآخر يميل الى التغير في اتجاه معين أيضاً . والمتغيران اما أن يتغيرا في نفس الاتجاه (بالزيادة أو بالنقص) فيسمى الارتباط طردياً ، واما أن يتغيرا في اتجاهين مختلفين فاذا اتجه أحدهما نحو الزيادة اتجه الآخر نحو النقص وفي هذه الحالة يسمى الارتباط عكسياً .

ووجود الارتباط لا يعني ان أي تغير في أحد المتغيرين لا بد أن يكون مصحوباً بتغير في المتغير الآخر ، وانما يعني أنه في أغلب الحالات يصحب التغير في احدي الظاهرتين تغير الظاهرة الأخرى ، وذلك هو السبب في اننا نقول أن التغير في أحد المتغيرين يصحبه ميل الظاهرة الاخرى نحو التغير . ويقاس الارتباط بمعامل الارتباط الذي يدلنا فقط على درجة العلاقة بين الظاهرتين ، هل هي علاقة قوية أو ضعيفة ، ولكنه لا يفسر لنا السبب في هذه العلاقة . والارتباط بين ظاهرتين يمكن أن يرجع الى واحد من الاسباب الآتية : -

١ - ان احد المتغيرين هو نتيجة مباشرة للمتغير الآخر مثل الارتباط بين سعر سلعة ما والطلب عليها أو عرضها .

٢ - ان احد المتغيرين هو سبب غير مباشر للمتغير الآخر ، فمثلا يكون ارتفاع ثمن القطن في مصر سبباً غير مباشر في ارتفاع ثمن القمح ، اذ يترتب على ارتفاع ثمن القطن زيادة المساحة المزروعة قطعاً فتقل المساحة المزروعة قمحاً وبذلك يقل عرضه في السوق فيرتفع ثمنه .

٣ - ان المتغيرين سوياً هما نتيجة لعامل واحد مشترك ، مثلاً التغير في سعر سلعتين تستهلكهما طبقة واحدة من الناس .

٤ - وجود عامل مشترك بين العوامل المختلفة التي تؤثر على كل من المتغيرين ، مثلاً الارتباط بين ذكاء الطالب في مادتين .

والمهم أن معامل الارتباط لا يحدد لنا أي من هذه الأسباب الأربعة هو سبب العلاقة التي ندرسها وإنما يدلنا فقط كما سبق ان ذكرنا عن شدة أو ضعف هذه العلاقة .

وقد يكون الغرض من حساب معامل الارتباط بجانب قياس درجة العلاقة بين ظاهرتين الاهتمام الى قيمة الظاهرة التابعة لوعرفنا قيمة الظاهرة المتبوعة . فلو عرفنا مثلاً ان هناك علاقة قوية بين سن العامل وأجرة الأمكننا ان نحدد الأجر الذي يجب ان يعطى لعامل في سن معين ، كما انه لو عرفنا ان هناك علاقة قوية بين الطول والوزن لأمكننا ان نحكم على ما يجب ان يكون عليه وزن شخص معين ذي طول محدد . وهذه العملية - استنباط قيمة متغير اذا عرفنا قيمة متغير آخر له علاقة به أساسها ما يسمى في الاحصاء بخطوط الانحدار وسوف نتعرض لهذا الموضوع فيما بعد . والمهم ان نعرف الآن ان استنباط قيمة ظاهرة اما اذا عرفنا قيمة ظاهرة أخرى لها علاقة بها يكون أمراً نظرياً فقط قد يتحقق فعلاً وقد لا يتحقق في الحياة العملية لأننا ذكرنا سابقاً ان الارتباط يعني ميل ظاهرة للتغير اذا تغيرت ظاهرة أخرى .

والارتباط اما ان يكون مستقيماً ويقاس بمعامل الارتباط واما ان يكون

غير مستقيم ويقاس بدليل الارتباط ونسبة الارتباط، فالأول هو ما لم يختلف نوعه من أول المدى الى آخره في الظاهرتين ، والثاني هو ما اختلف فيه طبيعة الارتباط خلال المدى فالارتباط بين مدة الخدمة في الحكومة والمهنية التي يتقاضاها الموظف ارتباط من نوع واحد من أول الى آخر التغير في الظاهرتين فهو لذلك ارتباط مستقيم . أما ارتباط بين كمية ما ينتجه عامل في مصنع في وحدة الزمن وبين طول وقت العمل ارتباط غير مستقيم فهو في بادىء الأمر ارتباط طردي حيث ان كمية المنتج تزيد كلما مكث العامل في عمله وقتاً أطول ثم تأتي اللحظة التي يصبح بعدها الارتباط سالباً حيث تنقص كمية المنتج كلما زاد وقت العمل .

وعند بحث الارتباط بين ظاهرة وأخرى نجد في كثير من الأحوال (الأمثلة)
العلاقة ظواهر أخرى ذات علاقات أخرى هامة بها وتؤثر على كل منها وعلى علاقتها (الأمثلة)
مع التوليد بعضها فلا يمكن صرف النظر عنها، وعندئذ إما أن نحسب معامل الارتباط
المتعدد للظاهرة التي ندرسها وحدها مع الظواهر الأخرى المؤثرة عليها مجتمعة.
أو أن نحسب معامل الارتباط الجزئي بين ظاهرة واحدة وظاهرة أخرى أو
أكثر مع حذف تأثير باقي الظواهر وذلك طبعاً حسب حاجة البحث.

جبريا بحيث لا تزيد قيمته عن الواحد الصحيح، فكلما قرب من الواحد الصحيح

معامل الارتباط جبرياً لا يزيد عن الواحد الموجب " هو لم
عكسي حيث ان الارتباط الجبري

كلما دل ذلك على قوة الارتباط . كما انه سمح لهذا المعامل بأن يأخذ إشارة
جبرية موجبة أو سالبة للدلالة على طبيعة الارتباط من حيث انه طردي أو
عكسي .

معامل بيرسون للارتباط :

نفترض أن لدينا ظاهرتان س ، ص وان كل منهما تتخذ القيم المتقابلة
الآتية :

الظاهرة	القيم
س ١	ص ١
س ٢	ص ٢
س ٣	ص ٣
س ٤	ص ٤
س ن	ص ن

وحيث أن مفهوم الارتباط هو مقارنة التغير الذي يحدث في الظاهرتين
فاننا نقيس هذا التغير بحساب انحراف كل قيمة لكل من الظاهرتين عن
وسطها الحسابي ، فاذا افترضنا أن الوسط الحسابي للظاهرة س هو س وان
الوسط الحسابي للظاهرة ص هو ص - يكون التغير المتقابل في الظاهرتين
هو كالآتي :

س ١ - س	ص ١ - ص
س ٢ - س	ص ٢ - ص
س ٣ - س	ص ٣ - ص
س ٤ - س	ص ٤ - ص
س ن - س	ص ن - ص

وحيث أن كل قيمة من هذه التغيرات هي قيمة مطلقة فلا يمكن بذلك مقارنة تغيرات الظاهرة س بتغيرات الظاهرة ص ولذلك نحول هذه التغيرات الى تغيرات نسبية بقسمة كل انحراف على الانحراف المعياري للظاهرة الخاصة به كالآتي :

$$\begin{array}{r} \frac{\text{ص} - \text{ص}_1}{\text{ص}} \\ \text{ع} \\ \text{ص} \\ \frac{\text{ص} - \text{ص}_2}{\text{ص}} \\ \text{ع} \\ \text{ص} \\ \frac{\text{ص} - \text{ص}_3}{\text{ص}} \\ \text{ع} \\ \text{ص} \\ \frac{\text{ص} - \text{ص}_4}{\text{ص}} \\ \text{ع} \\ \text{ص} \\ \frac{\text{ص} - \text{ص}}{\text{ن}} \\ \text{ع} \\ \text{ص} \end{array}$$

$$\begin{array}{r} \frac{\text{س} - \text{س}_1}{\text{س}} \\ \text{ع} \\ \text{س} \\ \frac{\text{س} - \text{س}_2}{\text{س}} \\ \text{ع} \\ \text{س} \\ \frac{\text{س} - \text{س}_3}{\text{س}} \\ \text{ع} \\ \text{س} \\ \frac{\text{س} - \text{س}_4}{\text{س}} \\ \text{ع} \\ \text{س} \\ \frac{\text{س} - \text{س}}{\text{ن}} \\ \text{ع} \\ \text{س} \end{array}$$

تقوم فكرة معامل بيرسون للارتباط على أساس انه اذا كان كل تغير في الظاهرة س يصحبه تغير في الظاهرة ص فان ضرب هذه الانحرافات يزيد من الناتج تطبيقاً لفكرة المغناطيسية . كذلك اذا كان كل تغير موجب في س يصحبه تغير موجب في ص فان ناتج الضرب يكون موجباً ويدل ذلك على ارتباط طردي - كذلك إذا كان كل تغير سالب في

إذا ضرب س₁ ص₁ ... و كان الناتج موجباً دل ذلك على ارتباط
 سواء كانت س₁ موجباً أو س₁ س₂ موجباً أو س₁ س₃ موجباً ...
 إذا كانت س₁ س₂ س₃ ... س_n س_{n+1} ... س_n س_{n+1} ...
 س_n يصبح تغير س₁ في ص₁ فان ناتج الضرب يكون موجباً ويدل ذلك
 أيضاً على ارتباط طردي . أما إذا كان التغير الموجب في س₁ يصحبه تغير
 س₁ في ص₁ أو العكس فان ناتج الضرب يكون سالب ويدل ذلك على
 ارتباط عكسي . وبذلك يرى بديسون ان ضرب الانحرافات س₁ عن وسطها
 الحسابي في الانحرافات ص₁ عن وسطها الحسابي يمكن أن يدلنا على قوة الارتباط
 وعلى نوعه إذا كان طردياً أو عكسياً (المقياس يسمى معامل الارتباط
 وأقصى حد له هو واحد صحيح سواء بالزائد أو بالنقص ويدل على وجود
 ارتباط كامل وكلما بعدت النتيجة عن واحد صحيح كلما قل الارتباط) .
 وحيث اننا في دراسة الارتباط لا يكون اهتمامنا موجهاً إلى كل حالة على
 حدة وانما إلى المتوسط العام لحاصل ضرب الانحرافات لذلك نحسب متوسط
 حاصل ضرب انحرافات س₁ في انحرافات ص₁ كالآتي :

$$\frac{1}{n} \left(\frac{s_1 - \bar{s}_1}{e} \times \frac{s_2 - \bar{s}_2}{e} \times \frac{s_3 - \bar{s}_3}{e} \times \dots \times \frac{s_n - \bar{s}_n}{e} \right)$$

$$\left(\frac{s_n - \bar{s}_n}{e} \times \frac{s_1 - \bar{s}_1}{e} \right)$$

$$(1) \quad \frac{(s_1 - \bar{s}_1)(s_2 - \bar{s}_2)}{e \times e} = \text{وبذلك يكون معامل الارتباط}$$

ومن الواضح انه يمكن وضع هذا القانون في شكل آخر كالآتي :

$$\frac{1}{n} \times (s - s) (s - s) = r$$

$$(2) \frac{\begin{matrix} ع \\ ص \end{matrix}}{\begin{matrix} ع \\ س \end{matrix}} = r$$

وبالمعالجة الجبرية للقيمة $\times (s - s) (s - s)$ نحصل على الآتي :

$$\text{حاصل ضرب انحرافات قيمة } s \text{ وقيمة } ص \text{ الاولى} =$$

$$s_1 s_1 - s_1 ص_1 - ص_1 s_1 + ص_1 ص_1$$

$$\text{حاصل ضرب انحرافات قيمة } s \text{ وقيمة } ص \text{ الثانية} =$$

$$s_2 s_2 - s_2 ص_2 - ص_2 s_2 + ص_2 ص_2$$

$$\text{حاصل ضرب انحرافات قيمة } s \text{ وقيمة } ص \text{ الثالثة} =$$

$$s_3 s_3 - s_3 ص_3 - ص_3 s_3 + ص_3 ص_3$$

$$\text{وبالجمع نحصل على } \times s - s - s - s + n s - s$$

وبالتعويض عن $\times s$ و $\times s$ نحصل على :

$$= \times s - n s - s - n s + n s - s$$

$$(3) \frac{\times s - n s - s - n s + n s - s}{\times s - n s - s - n s + n s - s} = r$$

$$\frac{\begin{matrix} \times s - n s - s - n s + n s - s \\ ع \\ ص \end{matrix}}{\begin{matrix} \times s - n s - s - n s + n s - s \\ ع \\ س \end{matrix}} = r$$

ويمكن وضع هذا القانون في صيغة أخرى كالتالي :

$$(4) \quad \frac{\text{مح س ص} - \text{س ص} - \text{س ص}}{\text{ع س} - \text{ع س}} = \text{ر}$$

واذا أخذنا انحرافات س عن وسط فرضي و وانحرافات ص عن وسط فرضي و نحصل على القانون .

$$\frac{\text{مح (س - و) (ص - و) - (ن ح س ح ص)}}{\text{ن ع س ع ص}} =$$

$$\begin{aligned} \text{حيث ح س} &= \text{س - و} \\ \text{ح ص} &= \text{ص - و} \end{aligned}$$

ويمكننا أن نثبت أن هذا القانون لا يختلف عن القوانين السابقة حيث أنه ليس الا اشتقاق جبري منها وذلك باجراء العمليات الجبرية الآتية :

حاصل ضرب انحرافات قيمة س وقيمة ص الاولى = س_١ ص_١ - س_١ و_١ - س_١ و_١ + و_١ -
 » » » » » الثانية = س_٢ ص_٢ - س_٢ و_٢ - س_٢ و_٢ + و_٢ -
 » » » » » الثالثة = س_٣ ص_٣ - س_٣ و_٣ - س_٣ و_٣ + و_٣ -

وبالجمع نحصل على مح س ص - ومح ص - و - مح س + ن و -

وبالتعويض عن مح س ، مح ص نحصل

$$\text{مح س ص} - \text{ن و ص} - \text{ن و} - \text{س} - \text{ن و} + \text{ن و} - \text{و}$$

والآن ننقل الى الجزء الثاني من القانون ن ح س ح ص وهو يساوي

$$\text{ن (س - و) (ص - و)}$$

$$= \text{ن س ص} - \text{ن س و} - \text{ن و ص} + \text{ن و و}$$

بطرح هذا الجزء من الجزء السابق طبقاً للقانون وبذلك نحصل على :

محس ص - ن و ص - ن و - س - + ن و و - - ن س - ص - +
 ن س - و - + ن و ص - - ن و و -

وهو يساوي محس ص - ن س - ص - وهو بسط القانون . رقم ٤ .
 وبذلك نلاحظ ان هذه القوانين ليست إلا صوراً مختلفة من القانون
 الاول ، إلا ان كلا منها يمكن أن يسهل علينا العمليات الحسابية التي تتفق
 مع أنواع البيانات الاحصائية التي تعرض علينا وسوف نرى ذلك في الامثلة
 التالية ، فإذا كانت انحرافات القيم عن أوساطها الحسابية ارقاماً صغيرة
 يكون استخدام القانون الاول أمراً بسيطاً من ناحية العمليات الحسابية التي
 يحتاجها ، أما إذا كانت القيم صغيرة ولكن انحرافاتهما عن أوساطها الحسابية
 انحرافات كبيرة (رقمياً) يكون استخدام القانون الثاني أسهل ، أما إذا
 كانت القيم كبيرة وانحرافاتهما عن أوساطها الحسابية قima كبيرة (رقمياً)
 يكون استخدام القانون الثالث أسهل . وفي التوزيعات التكرارية تكون
 العمليات الحسابية معقدة وطويلة إذا استخدمنا القانونين الاول والثاني ولذلك
 نقصر دائماً في هذه الحالة على استخدام القانون الثالث لأنه يبسط العمل الحسابي
 كثيراً . ولذلك تكون المفاضلة بين هذه القوانين هي فقط في حالة القيم غير
 المبوبة ، أما في حالة القيم المبوبة يجب أن يقتصر تفكيرنا على القانون الثالث .
 والمفاضلة هي فقط في ناحية سهولة أو صعوبة العمليات الحسابية ، إذ اننا
 نصل الى نفس النتيجة لمعامل الارتباط مهما كان القانون الذي نستخدمه من
 هذه القوانين الثلاث ، فهي كما قدمنا ليست إلا صوراً جبرية مختلفة لشيء
 واحد .

قانون سيرمان للارتباط :

وضع سيرمان قانوناً للارتباط على أساس ترتيب قيم كل من الظاهرتين
 س ، ص المتقابلتين في الشكل ر = $\frac{١ - ٦ \times ٢ \text{ ف}}{(١ - ٢ \text{ ن}) \text{ ن}}$ وقد استنبطه من

$$\text{القانون ر} = \frac{\text{مح س ص} - \text{ان س ص} - \text{كالاتي}}{\text{ن ع ع س ص}}$$

نفترض ان ترتيب قيم س هي ١، ٢، ٣، ٤، ن

وان ترتيب قيم ص المناظرة هي ٣، ٤، ١، ٢، ن وهي نفس الأرقام السابقة الخاصة بالظاهرة س ولكن في وضع مختلف .

$$\begin{aligned} \text{الوسط الحسابي لقيم س} &= \frac{\text{ن} + \text{١} + \text{٢} + \text{٣} + \text{٤} + \text{ن}}{\text{ن}} \\ &= \frac{\text{ن} (\text{١} + \text{ن})}{\text{ن}} \end{aligned}$$

(نظرية جبرية)

$$\frac{\text{ن} + \text{١}}{\text{٢}} = \frac{\text{ن} (\text{١} + \text{ن})}{\text{ن} \text{ ٢}} =$$

وحيث ان قيم ص هي نفس القيم وان كانت بوضع مختلف يكون الوسط الحسابي لها كذلك :

$$\frac{\text{ن} + \text{١}}{\text{١}}$$

أي ان س = ص -

$$\text{ع}^{\text{٢}} \text{ س} = \frac{\text{مح س ص}}{\text{ن}} - \text{س}^{\text{٢}}$$

$$\text{محد } ٢ = ١ + ٢ + ٣ + \dots + ٢٠٠٠ = \frac{(١ + ٢٠) (١ + ٢٠)}{٢} \quad (\text{نظرية جبرية})$$

$$\begin{aligned} \therefore \text{ع } ٢ \text{ س} &= \frac{(١ + ٢٠) (١ + ٢٠)}{٢} - \frac{(١ + ٢٠) (١ + ٢٠)}{٢} \\ &= \frac{١ + ٢٠ + ٢٠٠ + ٢٠٠٠}{٢} - \frac{(١ + ٢٠) (١ + ٢٠)}{٢} = \\ &= \frac{٣ - ٢٠ - ٢٠٠ - ٢ + ٢٠ + ٢٠٠}{٢} = \end{aligned}$$

$$\frac{١ - ٢٠}{٢} =$$

$$\frac{١ - ٢٠}{٢} = \text{ع } ٢ \text{ ص} \quad \text{وكذلك}$$

$$\text{محد } ٢ \text{ ف} = \text{محد } (٢ - ٢٠) \text{ ص}$$

$$= \text{محد } ٢ \text{ س} - ٢٠ \text{ محد } ٢ \text{ ص} + \text{محد } ٢ \text{ ص}$$

$$\frac{(١ + ٢٠) (١ + ٢٠)}{٢} = \text{وحيث أن محد } ٢ \text{ س} = \text{محد } ٢ \text{ ص} \text{ وكل منها}$$

$$\therefore \text{محد } ٢ \text{ ف} = \frac{(١ + ٢٠) (١ + ٢٠)}{٢} - ٢٠ \text{ محد } ٢ \text{ س} \text{ ص}$$

$$\therefore \text{محد } ٢ \text{ س} \text{ ص} = \frac{(١ + ٢٠) (١ + ٢٠)}{٢} - \frac{١}{٢} \text{ محد } ٢ \text{ ف}$$

$$\text{وحيث أن } r = \frac{\text{مح س ص} - \text{ن س} - \text{ص}}{\text{ن ع س ع ص}}$$

∴ بالتعويض عن مح س ص ، س- ، ص- ، ع س ، ع ص نحصل على :

$$r = \frac{\text{ن} \left(\frac{1 + \text{ن}}{2} \right) - \frac{1}{2} \text{مح ف}^2 - \frac{(1 + \text{ن}^2)(1 + \text{ن})}{6}}{\dots}$$

$$= r \cdot \frac{\text{ن} \left(\frac{(1 - \text{ن}^2)}{12} \times \frac{(1 - \text{ن}^2)}{12} \right)}{\frac{1}{2} \text{مح ف}^2 + \frac{\text{ن} + \text{ن}^2 + \text{ن}^3}{4} - \frac{\text{ن}^2 + 2\text{ن}^2 - \text{ن}^3}{6}}$$

$$= r \cdot \frac{\frac{1}{2} \text{مح ف}^2 + \frac{\text{ن} - \text{ن}^3}{12}}{\frac{\text{ن} - \text{ن}^3}{12} - \frac{\text{ن} - \text{ن}^3}{12}}$$

$$\therefore r = 1 - \frac{6 \text{ مح ف}^2}{\text{ن} (1 - \text{ن}^2)}$$

تطبيقات على قوانين الارتباط .

مثال ٥٣ :

من البيانات الآتية احسب الارتباط :

س	٦٥	٦٣	٦٧	٦٤	٦٨	٦٢	٧٠	٦٦	٦٨	٦٧	٦٩	٧١
ص	٦٨	٦٦	٦٨	٦٥	٦٩	٦٦	٦٨	٦٥	٧١	٦٧	٦٨	٧٠

٢ (س - س) (ص - ص)

ن ع ص

إذا أردنا تطبيق القانون الأول ر =

تكون خطوات العمل كالآتي :

ص	(س - س)	(ص - ص)	(س - س)	(ص - ص)	٢ (س - س)	٢ (ص - ص)
٦٨	١٧ -	٠٤	٠٦٨ -	٢٨٩	٠١٦	٣٨٩٢
٦٦	٣٧ -	١٦ -	٥٩٢	١٣١٩	٢٥٦	
٦٨	٠٣ +	٠٤ +	١٢	٠٩	٠١٦	
٦٥	٢٧ -	٢٦ -	٧٠٢	٧٢٩	٦٧٦	
٦٩	١٣ +	١٤ +	١٨٢	١٦٩	٥٩٦	
٦٦	٤٧ -	١٦ -	٧٥٣	٢٢٠٩	٢٥٦	
٦٨	٣٣ +	٠٤ +	١٣٢	١٠٨٩	٠١٦	
٦٥	٠٧ -	٢٦ -	١٨٢	٠٩	٦٧٦	
٧١	١٣	٣٤ +	٤٤٢	١٦٩	١١٥٦	
٦٧	٠٣ -	٠٦ -	٠١٨ -	٠٩	٠٣٦	
٦٨	٢٣ +	٠٤ +	٠٩٢	٥٢٩	٠١٦	
٧٠	٤٣ +	٢٤ +	١٠٣٢	١٨٤٩	٥٧٦	
٨١١	-	-	٤٠٣٤	٨٤٦٨	٣٨٩٢	

$$٦٧٦ = \frac{٨١١}{١٢} = -ص \quad ٦٦٧ = \frac{٨٠٠}{١٢} = -س$$

$$\frac{٤٠,٣٤}{٥٧٤٠} = \frac{٤٠,٣٤}{\frac{٣٨٩٢}{١٢} \times \frac{٨٤٦٨}{١٢} \sqrt{١٢}}$$

$$\therefore R = ٠,٧$$

وإذا أردنا تطبيق القانون الثاني $\frac{\text{محسب من ص - ن من ص}}{\text{ن ع س ص}} =$

تكون خطوات العمل كالآتي :

ص ^٢	س ^٢	س ص	ض	س
٤٦٢٤	٤٢٢٥	٤٤٢٠	٦٨	٦٥
٤٣٥٦	٣٩٦٩	٤١٥٨	٦٦	٦٣
٤٦٢٤	٤٤٨٩	٤٥٥٦	٦٨	٦٧
٤٢٢٥	٤٠٩٦	٤١٦٠	٦٥	٦٤
٤٧٦١	٤٦٢٤	٤٦٩٢	٦٩	٦٨
٤٣٥٦	٣٨٤٤	٤٠٩٢	٦٦	٦٢
٤٦٢٤	٣٩٠٠	٤٧٦٠	٦٨	٧٠
٣٢٢٥	٤٣٥٦	٤٢٩٠	٦٥	٦٦
٥٠٤١	٤٦٢٤	٤٨٢٨	٧١	٦١
٤٤٨٩	٤٤٨٩	٤٤٨٩	٦٧	٦٧
٤٦٢٤	٤٧٦١	٤٦٩٢	٦٨	٦٩
٤٩٠٠	٥٠٤١	٤٩٧٠	٨٠	٧١
٥٤٨٤٩	٥٣٤١٨	٥٤١٠٧	٨١١	٨٠٠

$$\frac{\frac{811}{12} \times \frac{800}{12} \times 12 - 54107}{\left(\left(\frac{811}{12} \right)^2 - \frac{54849}{12} \right) \times \left(\frac{800}{12} \right)^2 - \frac{53418}{12}} \sqrt{12} = r \therefore$$

$$0.7 = \frac{4034}{5740} = r \therefore$$

وإذا أردنا تطبيق القانون الثالث ر

بح (س - و) (ص - و) - د ح س ح ص

$$\frac{\text{ن ع ع}}{\text{س ص}}$$

تكون خطوات العمل كالآتي :

$$٠.٧ = \frac{٤٠.٣}{٥٧.٤} = \frac{٦٧ - ٤٧}{٥٧.٤} =$$

وإذا أردنا تطبيق قانون سيرمان ١ - $\frac{٦ \times ٢}{١ - ٢}$ تكون خطوات

العمل كالآتي :

(نلاحظ أن قيم كل من س ، ص ترتب ترتيباً تنازلياً وبذلك تكون
ف هي الفرق بين هذه التراتيب .

س	ص	تراتب س	تراتب ص	ف	٢ ف
٦٨	٦٨	٥	٥	٠.٥	٠.٢٥
٦٣	٦٦	١١	٩	١.٥	٢.٢٥
٦٧	٦٨	٧.٥	٥	٢	٤
٦٤	٦٥	١٠	١١	١.٥	٢.٢٥
٦٨	٦٩	٥	٣	٢	٤
٦٢	٦٦	١٢	٩	٢.٥	٦.٢٥
٧٠	٦٨	٢	٩	٣.٥	١٢.٢٥
٦٦	٦٥	٩	١١	٢.٥	٦.٢٥
٦٨	٧١	٥	١	٤	١٦
٦٧	٦٧	٧.٥	٨	٠.٥	٠.٢٥
٦٩	٦٨	٣	٥	٢.٥	٦.٢٥
٧١	٧٠	١	٢	١	١
-	-	-	-	-	٦١

(ملاحظة - اذا كانت القيم متشابهة تعطي متوسط التراتيب المسلسلة لها)

$$r = \frac{6 \times 6}{n(6-1)} - 1 = \dots$$

$$r = \frac{366}{1716} - 1 = \frac{61 \times 6}{(1 - 144) 12} - 1 =$$

$$r = 21 - 1 = 20$$

واذا كانت القيم مبوبة في توزيع تكراري مزدوج يكون من الافضل ،
كما قدمنا أن نقصر على استخدام القانون الثالث .

$$r = \frac{\text{معد (س - و) (ص - و) - ن ح س ح ص}}{\text{ن ع س ع ص}}$$

العمل كالآتي :

مثال ٥٤ :

الجموع	ص					س
	١٣ - ١١	١٢	١٣	١٤	١٥	
٣٨				٢٩	٩	صفر -
٢٤			١٢	١١	١	- ١
١٨		٥	١٣			- ٢
٢٠	٥	١٥				- ٣
١٠٠	٥	٢٠	٢٥	٤٠	١٠	الجموع

التوزيع الخاص بالظاهرة س :

ف	ك	م	ح	ح ك	ح ك
صفر -	٣٨	٥٠	١ -	٣٨ -	٣٨
١ -	٢٤	١٥	صفر	صفر	صفر
٢ -	١٨	٢٥	١ +	١٨ +	١٨
٣ - ٤	٢٠	٣٥	٢ +	٤٠ +	٨٠
المجموع	١٠٠	-	-	٢٠	١٣٦

$$س = -١٥ + ١ \times \frac{٢٠}{١٠٠} = ١٧$$

$$\sqrt{١٣٢} = \sqrt{١٣٦ - ٠.٠٤} = \sqrt{١٣٢} = ع$$

التوزيع الخاص بالظاهرة ص :

ف	ك	م	ح	ح ك	ح ك
٣ -	١٠	٤	٢ -	٢٠ -	٤٠
٥ -	٤٠	٦	١ -	٤٠ -	٤٠
٧ -	٢٥	٨	صفر	صفر	صفر
٩ -	٢٠	١٠	١ +	٢٠ +	٢٠
١١ - ١٣	٥	١٢	٢ +	١٠ +	٢٠
المجموع	١٠٠	-	-	٣٠ -	١٢٠

$$\text{ص}^- = 8 - \frac{30}{100} \times 2 = 7.4$$

$$\text{ع} = \sqrt{2 \left(\frac{30}{100} \right) - 1.2} = \sqrt{1.1} = 1.05$$

ولحساب χ^2 (س - و) (ص - و) نرجع إلى الجدول المزدوج ،
ومنه نكون الجدول الآتي الذي نضع فيه كل تكرار ورد في جسم الجدول
(ما عدا المجاميع) وأمامه انحرافه الواقعي الخاص بالظاهرة س
وانحرافه الواقعي الخاص بالظاهرة ص ونضرب الأرقام الثلاث في بعضها
كالتالي :

ك	س - و	ص - و	ك (س - و) (ص - و) (-)
٩	- ٤	- ١	٣٧
١	- ٤	صفر	صفر
٢٩	- ٢	- ١	٥٨
١١	- ٢	صفر	صفر
١٢	صفر	صفر	صفر
١٣	صفر	+ ١	صفر
٥	+ ٢	+ ١	١٠
١٥	+ ٢	+ ٢	٦٠
٥	+ ٤	+ ٢	٤٠
١٠٠	-	-	٣٠٤

ملاحظة :

يمكن إجراء هذه العملية على الجدول المزدوج نفسه حيث نضع أمام

فئات س وانحرافاتهما عن وسطها الفرضي (يجب ان تكون انحرافات واقعية غير مختصرة على طول الفئة) ، وأمام فئات ص انحرافاتهما عن وسطها الفرضي (يجب أن تكون انحرافات واقعية غير مختصرة على طول الفئة) .
 أمام الانحراف السيني صفر نشطب جميع الأرقام في جسم الجدول وكذلك أمام الانحراف الصادي صفر نشطب جميع الأرقام في جسم الجدول ، حيث ان هذه الأرقام سوف تضرب في صفر فيكون ناتجها = صفر . بعد ذلك نأخذ كل تكرار في جسم الجدول ونضربه في الانحراف السيني المقابل له في الانحراف الصادي المقابل له (لاحظ ان ثلاث أرقام تضرب في بعضها) ونضع النتيجة لدينا بإشارتها الجبرية ثم نجمع هذه النتائج جمعاً جبرياً .

على أساس النتائج السابق يكون معامل الارتباط ر كالآتي :

$$r = \frac{104 - 100 \times 2 \times 6}{132 \times 11 \sqrt{2 \times 1 \times 100}}$$

لاحظ ان ح س = س- و وحيث ان س- = و + $\frac{\text{م ح ك}}{\text{م ك}}$ × طول الفئة .

∴ س- - و = $\frac{\text{م ح ك}}{\text{م ك}}$ × طول الفئة ، وكذلك بالنسبة للظاهرة

حيث ح = ص- - و وحيث ان ص- = و- + $\frac{\text{م ح ك}}{\text{م ك}}$ × طول الفئة
 ∴ ص- - و- = $\frac{\text{م ح ك}}{\text{م ك}}$ × طول الفئة .

$$\therefore r = \frac{216}{2121} = 0.89$$

ويمكن اجراء العمل باستعمال نفس القانون ولكن بشكل آخر في الجدول الآتي :

معامل ارتباط الرتب للبيانات الوصفية

تصور ان احصائياً اجتماعياً قام بدراسة حالة سبع عائلات مختلفة في حي معين وسجل لكل عائلة الحالة العملية لرب الأسرة والمستوى الاقتصادي للأسرة نفسها فخرج بالجدول الآتي :

المستوى الاقتصادي للأسرة

الحالة العملية لرب الأسرة

فقيرة
معدمة
فقيرة
غنية
معدمة
متوسط الحال
فقيرة

يحمل شهادات متوسطة
أمي
يقرأ أو يكتب
يحمل شهادات عالية
أمي
أمي
يقرأ ويكتب

فاذا شاء هذا الاحصائي أن يقف على مدى العلاقة بين هاتين الظاهرتين يجب أن يستخدم معامل ارتباط الرتب وهو ما يسمى أحياناً بمعامل سيرمان نسبة الى صاحبه حيث ان البيانات المتوفرة لدينا بيانات وصفية ويمكن ترتيبها (لا يمكن استخدام معامل بيرسون حيث ان هذا المعامل لا يمكن تطبيقه الا في حالة البيانات القيمية) .

والفكرة الأساسية في قياس معامل ارتباط الرتب ، كما قدمنا ، هي مقارنة رتبتي الأسرة الواحدة في الظاهرتين فان اختلفتا كثيراً دل ذلك على قلة الارتباط وان اتحدا دل ذلك على شدة الارتباط أي ان أساس المعامل هو الفروق بين الرتب المتقابلة فكما كبرت هذه الفروق في المتوسط كلما ضعف الارتباط بين الظاهرتين والعكس كلما صغرت هذه الفروق .

ومعامل سيرمان الذي يقوم على هذا الأساس هو ١ - $\frac{6 \times 6}{n(n-1)}$

حيث $F^2 =$ مجموع مربعات الفروق بين الرتب المتقابلة

$n =$ عدد وحدات البحث .

ولحساب هذه المعامل نرتب الظاهرتين ترتيباً تدريجياً منتظماً ، وفي مثالنا السابق نعطي كل أسرة رتبة حسب الحالة العلمية لرب الأسرة والحالة الاقتصادية للأمره نفسها ثم نحسب الفرق بين الرتبتين ثم نربع هذا الفرق لزيادة حساسية المعامل ثم نطبق القاعدة السابقة - ويكون العمل كالآتي :

رتبة الحالة العملية	رتبة الحالة الاقتصادية		الفرق بين الرتب المعدلة		ف ^٢
	أولية	معدلة	أولية	معدلة	
٦	٦	(١)	٤	٢	٤٠٠
(١)	٢	(١)	١٥	٠.٥	٠.٢٥
(٤)	١٥	(٤)	٤	٠.٥	٠.٢٥
٧	٧	٧	٧	٠	٠٠
(٢)	٢	(٢)	١٥	٠.٥	٠.٢٥
(٣)	٢	٦	٦	٤	١٦٠٠
(٥)	٤.٥	(٥)	٤	٠.٥	٠.٢٥
—	—	—	—	—	٢١

$$r = (\text{معامل ارتباط الرتب}) = 1 - \frac{21 \times 6}{1 - 49 \times 7} = \frac{126}{336} - 1 =$$

$$= 375 - 1 = 374$$

(أي ان الارتباط طردي قوي نوعاً) .

نلاحظ ان الرتب المتشابهة وضمنها بين أقواس حسب ترتيبها الجاري في المجموعة ثم عدلنا هذه الرتب بأخذ المتوسط لها ، فالحالة العلمية يقرأ ريكيب تكررت مرتين وترتيبها هو ٤ ، ٥ ، فأخذنا المتوسط ٤.٥ كترتيب معدل . والحالة أمبي تكررت ثلاث مرات وترتيبها ١ ، ٢ ، ٣ ، فأخذنا المتوسط ٢ كترتيب معدل . وهكذا بالنسبة للظاهرة الثانية وهي الخاصة بالحالة الاقتصادية للأسرة نفسها .

وميزة هذا المعامل هي بساطة حسابه ، فعمليات الحسابية التي يحتاجها سهلة للغاية ولكن لا يجب أن نستخدمه في قياس الارتباط الا في حالة ما يكون ترتيب الظواهر أمراً منطقياً معقولاً .

معامل الارتباط للقيم المبوبة بطريقة الاقطار ذات الفروق المتساوية

او المجاميع المتساوية :

يمكن ايجاد معامل الارتباط من بيانات مبوبة بطريقة أسهل من طريقة بيرسون وتؤدي إلى نفس النتيجة تماماً ، وذلك بأن نرتب فئات من تصاعدياً أو تنازلياً ونرتب فئات من بنفس الكيفية ثم تتبع طريقة أقطار الفروق المتساوية أو طريقة أقطار المجاميع المتساوية . وتمتاز طريقة الرتب بأنها تصلح للاستعمال في الجداول التي تكون بها فئة أو أكثر من الفئات المفتوحة ولكنها لا تصلح مطلقاً للاستعمال في جدول تكون به الفئات غير متساوية .

وتعتمد هذه الطريقة على حساب الانحرافات المعيارية لرتب المتغيرين والفرق بين رتبتيها أو مجموعها . وتنحصر خطوات العمل في اعطاء كل فئة من فئات المتغيرين رتبة تبدأ بالواحد الصحيح وتنتهي بعدد مساو لعدد فئات المتغير . ثم رسم اقطار في الجدول المزدوج تمر من اليمين إلى اليسار ويمتاز كل قطر منها بتساوي الفرق بين الرتبة السينية والصادية في جميع أجزائه . ويحتاج حساب معامل الارتباط بعد ذلك لتطبيق القانون التالي :

$$r = \frac{\sum_{i=1}^n \sum_{j=1}^n \frac{e_i^2}{n} - \frac{e^2}{n}}{\sum_{i=1}^n \frac{e_i^2}{n} - \frac{e^2}{n}}$$

حيث $\sum_{i=1}^n e_i^2$ مربع الانحراف المعياري لرتب s

حيث $\sum_{i=1}^n e_i^2$ مربع الانحراف المعياري لرتب s

حيث $\sum_{i=1}^n e_i^2$ مربع الانحراف المعياري للفرق بين رتب s ورتب s وسيوضح استخدام هذا القانون من المثال الآتي :

مثال ٥٦ :

رتب س	رتب س	١	٢	٣	٤	٥	المجموع
١	أقل من ١	١	٢	٣	٤	٥	١٥
٢	١	١	٢	٣	٤	٥	١٥
٣	٢	١	٢	٣	٤	٥	١٥
٤	٣	١	٢	٣	٤	٥	١٥
٥	٤	١	٢	٣	٤	٥	١٥
٦	٥	١	٢	٣	٤	٥	١٥
٧	٦	١	٢	٣	٤	٥	١٥
٨	٧	١	٢	٣	٤	٥	١٥
٩	٨	١	٢	٣	٤	٥	١٥
١٠	٩	١	٢	٣	٤	٥	١٥
١١	١٠	١	٢	٣	٤	٥	١٥
١٢	١١	١	٢	٣	٤	٥	١٥
١٣	١٢	١	٢	٣	٤	٥	١٥
١٤	١٣	١	٢	٣	٤	٥	١٥
١٥	١٤	١	٢	٣	٤	٥	١٥
١٦	١٥	١	٢	٣	٤	٥	١٥
١٧	١٦	١	٢	٣	٤	٥	١٥
١٨	١٧	١	٢	٣	٤	٥	١٥
١٩	١٨	١	٢	٣	٤	٥	١٥
٢٠	١٩	١	٢	٣	٤	٥	١٥
٢١	٢٠	١	٢	٣	٤	٥	١٥
٢٢	٢١	١	٢	٣	٤	٥	١٥
٢٣	٢٢	١	٢	٣	٤	٥	١٥
٢٤	٢٣	١	٢	٣	٤	٥	١٥
٢٥	٢٤	١	٢	٣	٤	٥	١٥
٢٦	٢٥	١	٢	٣	٤	٥	١٥
٢٧	٢٦	١	٢	٣	٤	٥	١٥
٢٨	٢٧	١	٢	٣	٤	٥	١٥
٢٩	٢٨	١	٢	٣	٤	٥	١٥
٣٠	٢٩	١	٢	٣	٤	٥	١٥
٣١	٣٠	١	٢	٣	٤	٥	١٥
٣٢	٣١	١	٢	٣	٤	٥	١٥
٣٣	٣٢	١	٢	٣	٤	٥	١٥
٣٤	٣٣	١	٢	٣	٤	٥	١٥
٣٥	٣٤	١	٢	٣	٤	٥	١٥
٣٦	٣٥	١	٢	٣	٤	٥	١٥
٣٧	٣٦	١	٢	٣	٤	٥	١٥
٣٨	٣٧	١	٢	٣	٤	٥	١٥
٣٩	٣٨	١	٢	٣	٤	٥	١٥
٤٠	٣٩	١	٢	٣	٤	٥	١٥
٤١	٤٠	١	٢	٣	٤	٥	١٥
٤٢	٤١	١	٢	٣	٤	٥	١٥
٤٣	٤٢	١	٢	٣	٤	٥	١٥
٤٤	٤٣	١	٢	٣	٤	٥	١٥
٤٥	٤٤	١	٢	٣	٤	٥	١٥
٤٦	٤٥	١	٢	٣	٤	٥	١٥
٤٧	٤٦	١	٢	٣	٤	٥	١٥
٤٨	٤٧	١	٢	٣	٤	٥	١٥
٤٩	٤٨	١	٢	٣	٤	٥	١٥
٥٠	٤٩	١	٢	٣	٤	٥	١٥
٥١	٥٠	١	٢	٣	٤	٥	١٥
٥٢	٥١	١	٢	٣	٤	٥	١٥
٥٣	٥٢	١	٢	٣	٤	٥	١٥
٥٤	٥٣	١	٢	٣	٤	٥	١٥
٥٥	٥٤	١	٢	٣	٤	٥	١٥
٥٦	٥٥	١	٢	٣	٤	٥	١٥
٥٧	٥٦	١	٢	٣	٤	٥	١٥
٥٨	٥٧	١	٢	٣	٤	٥	١٥
٥٩	٥٨	١	٢	٣	٤	٥	١٥
٦٠	٥٩	١	٢	٣	٤	٥	١٥
٦١	٦٠	١	٢	٣	٤	٥	١٥
٦٢	٦١	١	٢	٣	٤	٥	١٥
٦٣	٦٢	١	٢	٣	٤	٥	١٥
٦٤	٦٣	١	٢	٣	٤	٥	١٥
٦٥	٦٤	١	٢	٣	٤	٥	١٥
٦٦	٦٥	١	٢	٣	٤	٥	١٥
٦٧	٦٦	١	٢	٣	٤	٥	١٥
٦٨	٦٧	١	٢	٣	٤	٥	١٥
٦٩	٦٨	١	٢	٣	٤	٥	١٥
٧٠	٦٩	١	٢	٣	٤	٥	١٥
٧١	٧٠	١	٢	٣	٤	٥	١٥
٧٢	٧١	١	٢	٣	٤	٥	١٥
٧٣	٧٢	١	٢	٣	٤	٥	١٥
٧٤	٧٣	١	٢	٣	٤	٥	١٥
٧٥	٧٤	١	٢	٣	٤	٥	١٥
٧٦	٧٥	١	٢	٣	٤	٥	١٥
٧٧	٧٦	١	٢	٣	٤	٥	١٥
٧٨	٧٧	١	٢	٣	٤	٥	١٥
٧٩	٧٨	١	٢	٣	٤	٥	١٥
٨٠	٧٩	١	٢	٣	٤	٥	١٥
٨١	٨٠	١	٢	٣	٤	٥	١٥
٨٢	٨١	١	٢	٣	٤	٥	١٥
٨٣	٨٢	١	٢	٣	٤	٥	١٥
٨٤	٨٣	١	٢	٣	٤	٥	١٥
٨٥	٨٤	١	٢	٣	٤	٥	١٥
٨٦	٨٥	١	٢	٣	٤	٥	١٥
٨٧	٨٦	١	٢	٣	٤	٥	١٥
٨٨	٨٧	١	٢	٣	٤	٥	١٥
٨٩	٨٨	١	٢	٣	٤	٥	١٥
٩٠	٨٩	١	٢	٣	٤	٥	١٥
٩١	٩٠	١	٢	٣	٤	٥	١٥
٩٢	٩١	١	٢	٣	٤	٥	١٥
٩٣	٩٢	١	٢	٣	٤	٥	١٥
٩٤	٩٣	١	٢	٣	٤	٥	١٥
٩٥	٩٤	١	٢	٣	٤	٥	١٥
٩٦	٩٥	١	٢	٣	٤	٥	١٥
٩٧	٩٦	١	٢	٣	٤	٥	١٥
٩٨	٩٧	١	٢	٣	٤	٥	١٥
٩٩	٩٨	١	٢	٣	٤	٥	١٥
١٠٠	٩٩	١	٢	٣	٤	٥	١٥

أولاً : حساب $\sum_{i=1}^n e_i^2$

س

رتب س	ك	ح	ح ك	ح ٢ ك
١	١٠	٢ -	٢٠ -	٤٠
٢	٤٠	١ -	٤٠ -	٤٠
٣	٢٥	صفر	صفر	صفر
٤	٢٠	١ +	٢٠	٢٠
٥	٥	٢ +	١٠	٢٠
المجموع	١٠٠	-	٦٠ -	
			٣٠ +	
			٣٠ -	

$$= \sqrt{120 - 100} = \sqrt{20} = 4.47$$

ثانياً : حساب ع^٢
ص

رتب ص	ك	ح	ح ك	ح ^٢ ك
١	٣٨	٢ -	٧٦ -	١٥٢
٢	٢٤	١ -	٢٤ -	٢٤
٣	١٨	صفر	صفر	صفر
٤	٢٠	١ +	٢٠	٢٠
المجموع	١٠٠	-	١٠٠ - ٢٠ + ٨٠ -	١٩٦

$$\sqrt{١٣٢} = \sqrt{٢ \left(\frac{٨٠}{١٠٠} \right) - \frac{١٩٦}{١٠٠}} = \sqrt{٢} \text{ ص}$$

ثالثاً : حساب ع^٢

ف

الفرق بين رتبتي س ، ص		ك	ك ف	ك ف ^٢
القطر	١		٥١	٥١
ب	صفر	٥١ = ٥ + ٥ + ١٢ + ٢٩	صفر	صفر
ج	١ -	٤٨ = ١٥ + ١٣ + ١١ + ٩	١ -	١
المجموع		١٠٠	٥٠	٥٢

$$\sqrt{r_{27}} = \sqrt{r_{25} - r_{52}} = \sqrt{\left(\frac{50}{100}\right) - \frac{52}{100}} = \sqrt{-2} = f_{28}$$

$$\frac{227 - 268}{12602 \sqrt{}} = \frac{227 - 132 + 111}{132 \times 111 \sqrt{2}} = \therefore$$

$$J_{19} = \frac{2516}{(1521.4) 2} =$$

$$\frac{\begin{array}{r} ٢٤ \\ - ٢٤ \\ \hline ٢٤ \end{array}}{\begin{array}{r} ٢٤ \\ - ٢٤ \\ \hline ٢٤ \end{array}} = \text{ويڪن استعمال قانون آخر هو ر}$$

حيث ع^٢ مربع الانحراف المعياري لرتب س

ع^٢ ص » » » » ص

ع ٢ » » » » المجموع رتب س، ص

ولتطبيق هذا القانون نرسم أقطاراً تمر من اليسار إلى اليمين ويمتاز كل قطر منها بتساوي مجموع رتب س ، ص في جميع أجزائه .

القطر	مجموع الرتب	ك	ح	ك ح	ك ح
أ	٢	٩	٣-	٢٧-	٨١
ب	٣	٣٠	٢-	٦٠-	١٢٠
ج	٤	١١	١-	١١-	١١
د	٥	١٢	صفر	صفر	صفر
هـ	٦	١٣	١+	١٣	١٢
و	٧	٥	٢+	١٠	٢٠
ل	٨	١٥	٣+	٤٥	١٣٥
م	٩	٥	٤+	٢٠	٨٠
المجموع	-	-	-	٩٨ - ٨٨ + ١٠ -	٤٦٠

$$ع^2 = \left(\frac{100}{100} \right)^2 - \frac{460}{100} \sqrt{}$$

$$= \sqrt{409} - \sqrt{46 - 401}$$

$$٨٩ = \frac{216}{(12104)^2} = \frac{111 - 132 - 409}{132 \times 111 \sqrt{2}} =$$

الارتباط بين الظواهر غير الرقمية :

نحتاج أحياناً إلى دراسة درجة العلاقة بين ظاهرتين لا يمكن التعبير عنها

بالارقام مثل الارتباط بين التطعيم بمصل ضد مرض معين والاصابة بهذا المرض أو بين الحالة التعليمية لمجموعة من الأشخاص ومستواهم المادي وغير ذلك من الأمثلة الكثيرة ، وفي هذه الحالة لا نستطيع أن نستخدم معامل الارتباط أو معامل الائتلاف لأن حساب هذين المعاملين يتوقف على معرفة الأوساط الحسابية وغير ذلك من المقاييس التي لا يمكن حسابها في حالة الظواهر غير الرقمية .

ولقياس الارتباط بين هذه الظواهر نستخدم معامل الاقتران والقانون الخاص به هو

$$r = \frac{a - d}{a + d + b + c}$$

ولشرح هذا القانون نأخذ مثلاً عملياً مع ملاحظة ان هذا القانون لا يمكن استخدامه إلا في الحالات التي تنقسم فيها الظواهر التي ندرس العلاقة بينها إلى نوعين فقط ، أما فيما عدا ذلك من الحالات يجب استخدام مقياس آخر سيأتي الكلام عنه فيما بعد . فإذا فرضنا انه أثناء وباء للتيفود مثلاً أجرى أحد الأطباء تجربة مصل جديد على عينة من الأفراد عددها ن ، وقد طعم منهم بالمصل عدداً معيناً فأصيب بالمرض العدد د ولم يصب العدد ا - أما الباقي الذين لم يطعموا فأصيب منهم العدد د ولم يصب العدد ب ، فكيف يمكن من هذه البيانات أن ندرس درجة العلاقة بين الحقن بالمصل وعدم الاصابة .

يمكننا وضع البيانات السابقة في الجدول التالي .

مثال ٥٧ :

المجموع	لم يطعموا	طعموا	
٣٠٥	ب ١١٣	ا ١٩٢	لم يصب
٣٨	د ٣٤	ح ٤	أصيب
٣٤٣	١١٧	٢٢٦	المجموع

$$\therefore \text{معامل الاقتران} = \frac{٤ \times ١١٣ - ٣٤ \times ١٩٢}{٤ \times ١١٣ + ٣٤ \times ١٩٢} = ٨٧ =$$

وتدل هذه النتيجة على وجود علاقة طردية قوية بين التطعيم بالمصل وعدم الإصابة بالمرض . وقيمة هذا المعامل تكون دائماً أقل من ١ طالما كانت ب > أكثر من صفر ، أما إذا كانت تساوي صفر فإن المعامل = ١ وفي هذه الحالة تكون هناك علاقة كاملة بين الظاهرتين التي نقوم ببجتها .

ونتيجة المثال السابق يمكن تحقيقها إذا حسبنا نسبة الذين طعموا ولم يصابوا بالمرض وهي ٩٨ ٪ تقريباً ، بينما نسبة الذين لم يطعموا ولم يصابوا تساوي ٧٦ ٪ تقريباً .

أما إذا كانت إحدى الظاهرتين اللتين نبجث العلاقة بينهما أو كليهما تنقسم إلى أكثر من نوعين ، فإن معامل الاقتران لا يساعدنا في هذه الحالة وعندئذ

نستخدم معامل التوافق الذي وضعه بيرسون لقياس العلاقة بين الصفات غير المقيسة ، أو بين صفات بعضها يقاس بالأرقام وبعضها لا يقاس .

ويحسب معامل التوافق من القانون التالي :

$$\text{معامل التوافق} = \sqrt{\frac{1 - \chi^2}{n}}$$

ولتطبيق هذا القانون نبوب البيانات الخاصة بالظاهرتين في جدول مزدوج ثم نربع كل تكرار في الجدول ونقسمه على حاصل ضرب التكرار الكلي العمودي في التكرار الكلي الأفقي .

مثال ٥٨ :

البيانات التالية خاصة بعدد الطلبة حسب تقديرات نجاحهم في مادتين :

م	مقبول	جيد	جيد جدا	ممتاز	المجموع
مقبول	٢٧	٥٠	٣٠	٢١	١٢٨
جيد	٤٥	١٠٣	٨٢	٥٠	٢٨٠
جيد جدا	١٠	٤٨	٢٣	١٤	٦٥
المجموع	٨٢	١٧١	١٣٥	٨٥	٤٧٣

السطر الأول : $\frac{1}{128} = \left(\frac{27}{82} + \frac{50}{171} + \frac{30}{135} + \frac{21}{85} \right) \times 270.2$

السطر الثاني : $\frac{1}{280} = \left(\frac{45}{82} + \frac{103}{171} + \frac{82}{135} + \frac{50}{85} \right) \times 927$

السطر الثالث : $\frac{1}{65} = \left(\frac{10}{82} + \frac{48}{171} + \frac{23}{135} + \frac{14}{85} \right) \times 437$

$$\chi^2 = 270.2 + 927 + 437 = 1674.2$$

$$\text{معامل التوافق} = \sqrt{\frac{1 - 1674.2}{1674.2}}$$

ونلاحظ هنا النتيجة على عدم وجود علاقة بين هاتين الظاهرتين

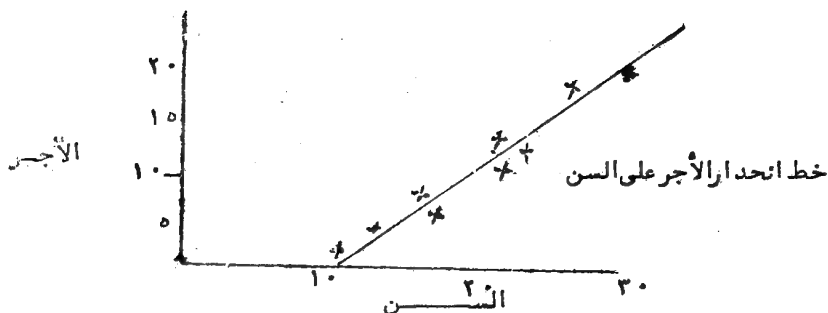
معظمها عليه فان هذا دليل على شدة الارتباط بين الظاهرتين بينما إذا بعدت معظم النقط عن خط الانحدار فان هذا دليل على ضعف الارتباط بينهما .
ويعنى آخر كلما كان تشتت النقط حول خط الانحدار كبيراً كلما ضعف الارتباط والعكس كلما كان تشتتها ضعيفاً كان هذا دليل على شدة الارتباط بين الظاهرتين التي ندرس العلاقة بينهما .

وخط الانحدار بذلك يوضح لنا هندسياً العلاقة بين الظاهرتين ، فهو يرينا كيف تميل الظاهرة التابعة إلى التغير نتيجة تغير معين في الظاهرة المستقلة .
على أن هذا التغير في الظاهرة التابعة الذي يظهره الرسم ليس هو حتماً نفس التغير الذي يحدث في الواقع العملي ، تماماً مثل المتوسط لعدة قيم ، فلا يعنى هذا المتوسط أن جميع القيم متساوية حتماً ولذا فخط الانحدار يسمى أحياناً خط العلاقة المتوسطة بين الظاهرتين حيث يعطينا القيمة النظرية للمتغير التابع التي تقابل قيمة معينة للمتغير المستقل . هذه القيمة الثالثة قد تكون هي نفسها القيمة الواقعية وقد تختلف عنها بعض الشيء . وكلما قربت القيم النظرية للمتغير التابع من قيمته الواقعية كلما كان ذلك دليلاً على أن خط الانحدار يمثل العلاقة بين الظاهرتين تمثيلاً صادقاً ، وكلما بعدت القيم النظرية عن القيم الواقعية كلما دل ذلك على ضعف تمثيل خط الانحدار للعلاقة بين الظاهرتين .

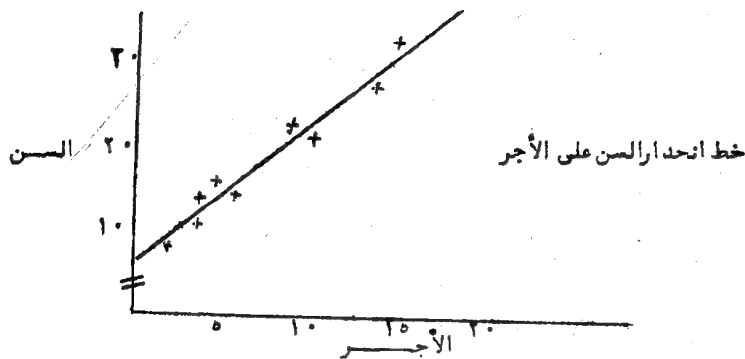
الأجر اليومي بالقرش	سن الطفل
٢	١٢
٣	١٤
٩	٢١
٦	١٦
١١	٢٠
١٦	٢٥
١٩	٣٠
٣	١٥
٤	١٧
٧	٢٠

لو مثلنا سن العامل على المحور الأفقي وأجره على المحور الرأسي ومثلنا كل زوج من القيم المتقابلة في الجدول السابق بنقطة بيانية لانتشرت النقاط بنظام خاص يسمى بالاتجاه العام لها ، ولذل ذلك على نوع العلاقة بين الظاهرتين، ويمكن تمثيل هذه العلاقة بخط مستقيم كما فعلنا في الشكل. ويسمى الشكل الذي نحصل عليه باتباع هذه الخطوات في الرسم بشكل الانتشار (Scatter Diagram) والمستقيم الممثل له بخط الانحدار أو خط العلاقة المتوسطة .

وخط الانحدار السابق هو خط انحدار الأجر على السن حيث انه جرت العادة على وصف الخط باسم الظاهرة التابعة أولاً ويتضح للقارىء سبب هذه التسمية عند الكلام على فائدة خطوط الانحدار .



ويمكننا كذلك ان نعتبر ان الأجر هو المتغير المستقل فنرسمه على المحور الافقي وان السن هو المتغير التابع فنرسمه على المحور الرأسي وبذلك نحصل على خط انحدار السن على الأجر كما يتضح من الرسم التالي :



وفائدة خطوط الانحدار انها تمكن الباحث من تقدير قيمة المتغير التابع لو عرف قيمة المتغير المستقل ؛ فلو قيل لك ما الأجر الذي يجب أن يعطى لعامل عمره ٢٢ سنة مثلا ، لأمكن استخدام الشكل الأول في استنباط القيمة

باقامة عمود على المحور الافقي عند السن ٢٢ سنة فيقابل خط الانحدار في نقطة يكون احداثيها الراسي هو القيمة المطلوبة (١١ر٠٩ قرشاً) . وبالمثل يمكن استخدام خط انحدار السن على الأجر لتحديد ما نتوقع أن يكون عليه عمر العامل ذي أجر معين .

وتحديد القيم التابعة على أساس خطوط الانحدار بطريقة بيانية تعوزها الدقة وتعتمد الى حد كبير على الحكم الشخصي للباحث نفسه حيث عليه يتوقف رسم خط الانحدار وهو أمر ليس من السهل ولا يمكن ان يتفق فيه اثنان الا نادرا . لهذا يجب تمثيل العلاقة بطريقة جبرية دقيقة .

طرق ايجاد معادلة الانحدار :

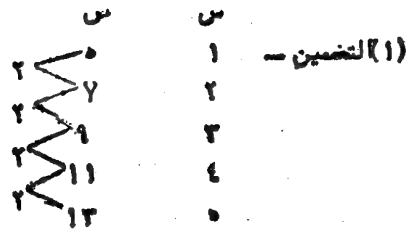
لايجاد معادلة الانحدار يجب أولاً رسم شكل الانتشار حتى نتمكن ما اذا كانت العلاقة بين الظاهرتين من الدرجة الاولى اي يصورها خط مستقيم أو من درجة اعلى أي يصورها خط غير مستقيم .

فاذا كانت العلاقة من الدرجة الاولى فنعمل على ايجاد معادلة المستقيم من واقع القيم المتناظرة التي تتخذها الظاهرتين ، وذلك إما بطريقة النضمين إذا تبين لنا ان جميع النقط التي تمثل القيم المتناظرة تقع على خط مستقيم ، أما إذا كانت النقط تنتشر حول خط مستقيم فيمكن ايجاد المعادلة اما بطريقة أصغر المربعات وإما بطريقة تعتمد على كون العلاقة بين الظاهرتين تتوقف على درجة تشتت النقط حول خط العلاقة المتوسطة (خط الانحدار) .

مثال ٦٠ :

ايجاد معادلة خط انحدار $\frac{ص}{س}$ من واقع القيم الآتية :

(استخدمنا طريقة التضمين حيث
إذا رسمنا شكل الانتشار تفتح
جميع النقاط على خط مستقيم)



$$ص = ٥ + ٢ (س - ١)$$

$$٥ + ٢س - ٢$$

$$= ٢س + ٣ \text{ وهي المعادلة المطلوبة .}$$

(٢) بتطبيق طريقة أصغر المربعات

$$ص = م + س + ب$$

$$مح ص = م مح س + ن ب \quad (١)$$

$$مح س ص = م مح س + ب مح س \quad (٢)$$

وبالتعويض في هاتين المعادلتين يمكن أن نحسب قيمة كل من م ، ب

س	ص	س	س ص
١	٥	١	٥
٢	٧	٤	١٤
٣	٩	٩	٢٧
٤	١١	١٦	٢٤
٥	١٣	٢٥	٦٥
١٥	٤٥	٥٥	١٥٥

$$٤٥ = ١٥م + ٥ب \quad (٣)$$

$$١٥٥ = ٥٥م + ١٥ب \quad (٤)$$

وبضرب المعادلة -١ في ٣ تكون

$$١٣٥ = ٤٥م + ١٥ب$$

$$١٥٥ = ٥٥م + ١٥ب$$

$$\text{وبالطرح } ٢٠ = ١٠م$$

$$٢ = م$$

وبالتعويض في المعادلة رقم ٣ - يكون :

$$٤٥ = ٣٠ + ٥ ب$$

$$١٥ = ٥ ب$$

$$٣ = ب$$

ص = ٢ س + ٣ وهي المعادلة المطلوبة .

وبالمثل يمكن إيجاد معادلة س على ص حيث نستخدم الصيغة س = م - ص + ب .

(٣) على أساس معرفة الوسط الحسابي لكل من الظاهرتين وانحرافهما المعياري ودرجة الارتباط بينهما (معامل الارتباط) يمكن إيجاد معادلة الانحدار كالتالي :

$$\text{معادلة الانحدار } \frac{\text{ص}}{\text{س}} \text{ هي (ص - ص -) } = \text{ر} \frac{\text{ع ص}}{\text{ع س}} \text{ (س - س -)}$$

$$\text{ومعادلة الانحدار } \frac{\text{س}}{\text{ص}} \text{ هي (س - س -) } = \text{ر} \frac{\text{ع س}}{\text{ع ص}} \text{ (ص - ص -)}$$

$$\text{ومن هاتين المعادلتين نعرف ان معامل الانحدار } \frac{\text{ص}}{\text{س}} = \text{ر} \frac{\text{ع ص}}{\text{ع س}}$$

وهو يساوي ظل زاوية انحدار خط الانحدار $\frac{\text{ص}}{\text{س}}$ ومعامل الانحدار $\frac{\text{س}}{\text{ص}}$

$$= \text{ر} \frac{\text{ع س}}{\text{ع ص}} \text{ وهو يساوي ظل زاوية انحدار خط الانحدار } \frac{\text{س}}{\text{ص}} . \text{ فادا علم لدينا}$$

معامل الانحدار $\frac{\text{ص}}{\text{س}}$ ومعامل انحدار $\frac{\text{س}}{\text{ص}}$ نستطيع بضرهما وإيجاد الجذر

للتائج أن نحصل على معامل الارتباط حيث $r = \frac{ع ص}{ع س} \times \frac{ر ع}{ر ص} = r^2$

فإذا رجعنا للقيم السابقة امكننا إيجاد معادلة انحدار $\frac{ص}{س}$ كالآتي :

س	ص	س ²	ص ²
١	٥	١	٢٥
٢	٧	٤	٤٩
٣	٩	٩	٨١
٤	١١	١٦	١٢١
٥	١٣	٢٥	١٦٩
<u>١٥</u>	<u>٤٥</u>	<u>٥٥</u>	<u>٤٤٥</u>

$$س = \frac{١٥}{٥} = ٣ ، ص = \frac{٤٥}{٥} = ٩$$

$$ع س = \sqrt{٩ - \frac{٥٥}{٩}} = \sqrt{٢}$$

$$ع ص = \sqrt{٨١ - \frac{٤٤٥}{٩}} = \sqrt{٨١ - ٨٩} = \sqrt{٨}$$

ان جميع النقط تقع على خط مستقيم يكون الارتباط كامل كما أشرنا إلى ذلك من قبل .

وبذلك تكون معادلة انحدار $\frac{ص}{س}$ هي

$$(3 - 5) \cdot \frac{\sqrt{1}}{\sqrt{2}} \times 1 = 9 - 5$$

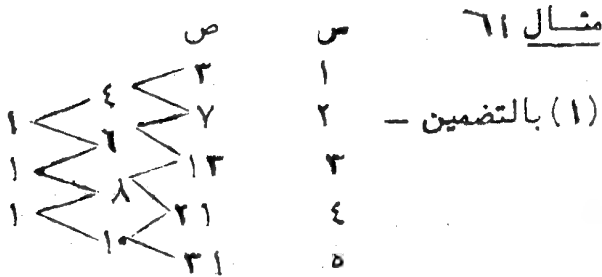
∴ ص - ۹ = ۲ س - ۶

∴ ص = ٢ س + ٣ وهي المعادلة المطلوبة

وبالمثل يمكن إيجاد معادلة انحدار $\frac{س}{ص}$

طرق إيجاد معادلة الانحدار (من الدرجة الثانية)

إذا تبين من شكل الانتشار ان المعادلة من الدرجة الثانية أي ان خط الانحدار يكون منحنى نستطيع إيجاد معادلة الانحدار اما بطريقة التضمين إذا كانت جميع النقاط تقع على المنحنى أو بطريقة أصغر المربعات إذا كانت تنتشر حوله .



$$1 \text{ ص} = 2 + 4(1 - \text{س}) + 1(1 - \text{س}) + 2(1 - \text{س})$$

$$2 + 3s - 2 + 4 - 4s =$$

وهي المعادلة المطلوبة $\bar{S} = S + 1$

(٢) بطريقة أصغر المربعات

المعادلة العامة للمنحنى

$$ص = ا س^٢ + ب س + ح$$

١ - : مع ص = ا مع س^٢ + ب مع س + ح

٢ - ' مع س = ا مع س^٣ + ب مع س^٢ + ح مع س

٣ - ' مع س^٢ = ا مع س^٤ + ب مع س^٣ + ح مع س^٢

من هذه المعادلات الثلاث نستطيع ان نحسب قيمة ا ، ب ، ح

س	ص	ص	س ^٢	س ^٣	س ^٤	س ^٥
١	٣	٣	١	١	١	٣
٢	٧	١٤	٤	٨	١٦	٣٨
٣	١٣	٣٩	٩	٢٧	٨١	١١٧
٤	٢١	٨٤	١٦	٦٤	٢٥٦	٢٣٦
٥	٣١	١٥٥	٢٥	١٢٥	٦٢٥	٧٧٥

$$١٠ \quad ٧٥ \quad ٢٩٥ \quad ٥٥ \quad ٢٢٥ \quad ٩٧٩ \quad ١٢٥٩$$

١ - : ٧٥ = ٥٥ + ا ١٥ + ب ٥

٢ - ' ٢٧٥ = ٥٥ + ا ٢٢٥ + ب ١٥

٣ - ' ١٢٥٩ = ٩٧٩ + ا ٢٢٥ + ب ٥٥

بضرب المعادلة - ١ - في ٣ ينتج :

$$٢٢٥ = ١٦٥ + ا ٤٥ + ب ١٥$$

$$٢٩٥ = ٢٢٥ + ا ٥٥ + ب ١٥$$

٤ - بالطرح ينتج ٧٠ = ٦٠ + ا ١٠ + ب

وبضرب المعادلة - ١ - في ٣ ينتج

$$٨٢٥ = ١٦٠ + ب + ٥٥$$

$$١٢٥٩ = ١٩٧٩ + ب + ٥٥$$

$$\text{بالطرح ينتج } ٤٣٤ = ١٤٣٤ + ب$$

بضرب المعادلة ٤ في ٦ ينتج :

$$٤٢٠ = ٣٦٠ + ب$$

$$٤٣٤ = ٣٧٤ + ب$$

$$\text{بالطرح ينتج } ١٤ = ١٤$$

$$١ = أ :$$

وبالتعويض في المعادلة ٤ عن قيمة أ ينتج :

$$٧٠ = ١٠ + ب$$

$$١٠ = ١٠ + ب$$

$$١ = ب :$$

وبالتعويض في المعادلة - ١ - عن قيمة أ ، ب ينتج :

$$٧٥ = ٥ + ١٥ + ٥٥$$

$$٧٥ = ٥ + ٧٠ :$$

$$٥ = ٥ :$$

$$١ = ١ :$$

المعادلة هي $ص = س^٢ + س + ١$ وهي المعادلة المطلوبة .

ايجاد معادلة الانحدار في حالة القيم المبوبة

لا تختلف خطوات العمل عند ايجاد معادلة الانحدار في حالة القيم المبوبة عن الخطوات السابقة الا في أنه يجب أولاً تحويل البيانات الواردة في الجدول المزدوج الى قيم بسيطة لكل من الظاهرتين . وذلك بأن نأخذ مراكز فئات المتغير المستقل ونعتبرها قيماً لهذا المتغير ثم نبحت عن القيم التابعة بحساب المتوسط المقابل لكل قيمة من قيم المتغير المستقل من واقع الجدول المزدوج .

بعد ذلك يصبح لدينا قيم بسيطة لكل من المتغيرين فنزسم شكل الانتشار ومنه نحدد درجة المعادلة ، فاذا كان الاتجاه العام مستقيماً عملنا على ايجاد معادلة المستقيم $y = mx + c$ ب بالطرق التي سبق مناقشتها واذا كان منحنى عملنا على ايجاد معادله الدرجة الثانية $y = ax^2 + bx + c$

مثال ٦٢ :

الآتي بيان بتوزيع تكراري مزدوج لايحار ١٠٠ موظف والمرتبات التي يتقاضونها . المطلوب ايجاد معادلة انحدار الايحار على المرتب .

المرتب الايحار	-٣	-٥	-٧	-٩	-١١	المجموع
صفر -	٩	٢٩				١٨
- ١	١	١١	١٢			٢٤
- ٢			١٣	٥		١٨
- ٣				١٥	٥	٢٠
المجموع	١٠	٤٠	٢٥	٢٠	٥	١٠٠

الحل :

خط انحدار ايحار المسكن على المرتب هو الخط البياني الذي يمثل العلاقة

بين مراكز فئات المراتب (المتغير المستقل) ومتوسط الايجار الذي يدفعه أفراد كل فئة (المتغير التابع) ويمكن حساب هذا المتوسط بضرب مركز كل فئة لاييجار المسكن في التكرار المقابل له في الحانة المواجهة لمركز فئة المتغير المستقل .

فمتوسط الايجار المقابل للمرتب ٤ (مركز الفئة ٣ - ٥) =

$$٦ = \frac{١ \times ١٥٥ + ٩ \times ٥}{١٠}$$

ومتوسط الايجار المقابل للمرتب ٦ (مقابل الفئة ٥ - ٧) =

$$٧٨ = \frac{١١ \times ١٥٥ + ٢٩ \times ٥}{٤٠}$$

ومتوسط الايجار المقابل للمرتب ٨ (مركز الفئة ٧ - ٩) =

$$٢٠٢ = \frac{١٣ \times ٢٥٥ + ١٣ \times ١٥}{٢٥}$$

ومتوسط الايجار المقابل للمرتب ١٠ (مركز الفئة ٩ - ١١) =

$$٢٥ = \frac{١٥ \times ٣٩٩ + ٥ \times ٢٥}{٢٠}$$

ومتوسط الايجار المقابل للمرتب ١٢ (مركز الفئة ١١ - ١٣) =

$$٣٥ = \frac{٥ \times ٣٥}{٥}$$

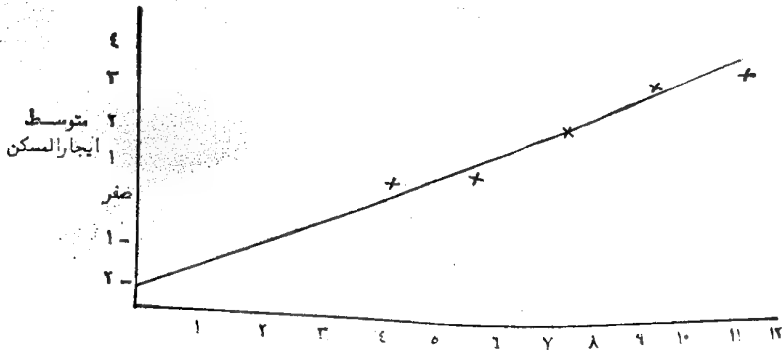
وبذلك يمكن وضع قيم المتغيرين في الشكل الآتي :

متوسط ايجار المسكن

المرتب

٦ر	٤
٧٨ر	٦
٢٠٢ر	٨
٣٢٥ر	١٠
٣٥ر	١٢

بعد ذلك نرمم هذه البيانات في شكل بياني كالآتي :



المرتب

من شكل الانتشار يتبين ان الاتجاه العام للعلاقة بين هذين المتغيرين هو مستقيم فالمعادلة التي تصور العلاقة بينها تكون من الدرجة الاولى . وحيث ان جميع النقط لا تقع على خط الانحدار لهذا لا يمكن ايجاد المعادلة بطريقة التضمين . وبذلك لا يمكن ايجاد المعادلة الا بطريقة اصغر المربعات أو

باستخدام الصيغة $ص - ص = ر \frac{عص}{عس} (س - س) .$ وخط انحدار

المرتب على ايجار المسكن هو الخط البياني الذي يمثل العلاقة بين مراكز فئات الايجار (المتغير المستقل) ومتوسطات المرتب المقابلة لهذه المراكز .

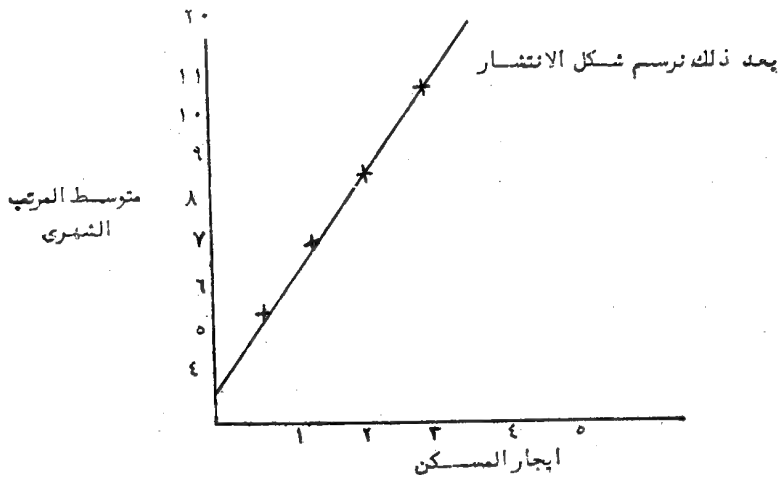
ومن الجدول السابق يمكن إيجاد متوسطات المراتب المقابلة لمراكز فئات
الايجار كالاتي :

$$\text{متوسط المراتب المقابل للايجار ٥ر (مركز الفئة صفر - ١)} = \frac{٢٩ \times ٦ + ٩ \times ٤}{٣٨} = ٥٥٣$$

$$\text{متوسط المراتب المقابل للايجار ١ر (مركز الفئة ١ - ٢)} = \frac{١٢ \times ٨ + ١١ \times ٦ + ١ \times ٤}{٢٤} = ٩٢ر$$

$$\text{» » » (مركز الفئة ٢ - ٣)} = \frac{٥ = ١٠ + ١٣ \times ٨}{١٨} = ٨٥٥ر$$

$$\text{» » » (مركز الفئة ٣ - ٤)} = \frac{٥ \times ١٢ + ١٥ \times ١٠}{٢٠} = ١٠٥ر$$



من شكل الانتشار يتبين أن المعادلة من الدرجة الأولى حيث يظهر أن الاتجاه العام للعلاقة بين المتغيرين مستقيم . ولا يمكن استعمال طريقة التضمين لايحاد لمعادلة لأن جميع النقط ليست على خط الاتجاه العام ، ولذلك نستعمل اما طريقة أصغر المربعات أو الطريقة التي تعتمد على الصيغة :

$$ص - ص = -ص - ر \frac{عص}{عس} (س - س -)$$

ايحاد معامل الارتباط بمعلومية الانحدار :

يتبين لنا مما سبق أن معامل الارتباط =

$$\frac{\text{معامل انحدار ص}}{\text{معامل انحدار س}} \times \frac{\text{معامل انحدار ص}}{\text{معامل انحدار س}}$$

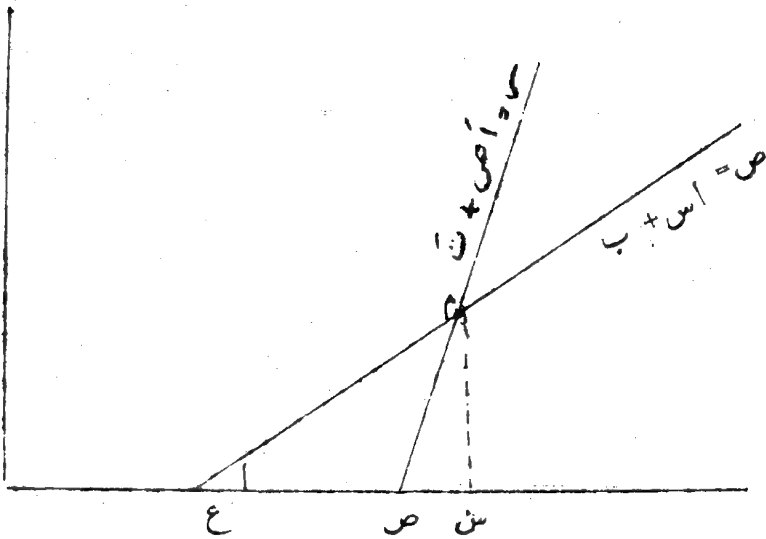
ولهذا اذا علم لدينا كل من هذين المعاملين عن طريق معرفة معادلتى الانحدار $ص = م س + ب$ ، $ص = م - ص + ب$ يمكن حساب معامل الارتباط بايحاد الجذر التربيعي لمعاملي الانحدار وهما م ، م- في هاتين المعادلتين .

$$\text{وبذلك } ر = \sqrt{م \times م -}$$

وقد يوجد لدينا خطي الانحدار فيمكن حساب معاملي الانحدار عن طريق حساب ميل كل خط من الخطين . وميل خط الانحدار هو ظل زاوية الانحدار ، ويساوي $\frac{\text{المقابل}}{\text{المجاور}}$.

مثال ٦٣ :

من الرسم التالي الذي يظهر فيه خطي الانحدار ، احسب معامل الارتباط :



$$\begin{aligned} \text{ميل خط الانحدار } ص = م + ب \\ \text{ظل زاوية الانحدار} \end{aligned} \quad \begin{aligned} \frac{م}{ص} = م ع س \\ \text{ع س} \end{aligned}$$

$$\begin{aligned} \text{ميل خط الانحدار } م = ص + ن \\ \text{ظل زاوية الانحدار} \end{aligned} \quad \begin{aligned} \frac{ص}{م} = ص م س \\ م س \end{aligned}$$

$$\frac{م}{ص} \sqrt{\frac{ص}{م}} = \frac{ص}{م} \sqrt{\frac{م}{ص}} \times \frac{م}{ص} \sqrt{\frac{ص}{م}} = ر \therefore$$

الخطا المعياري :

ذكرت فيما سبق ان العلاقة بين المتغيرين يمثلها خط الانحدار الذي يمكن تسميته بخط العلاقة المتوسطة . كذلك ذكرت ان النقط في شكل الانتشار

لا تقع جميعها في أغلب الأحيان على خط الانحدار وإنما تنتشر حوله ، وإن
خط العلاقة المتوسطة يمكن تشبيهه بالوسط الحسابي لمدّة قيم . فالوسط الحسابي
يمثل مجموعة القيم ويكون مجموع مربع انحرافات هذه القيم عنه أقل من
مجموع مربع انحرافات القيم عن أي قيمة أخرى غيره (كذلك خط العلاقة
المتوسطة فهو يمثل النقط المنتشرة حوله على أساس أن مجموع مربع انحرافات
النقط عنه أقل من مجموع مربع انحرافات عن أي خط آخر) وهذا هو
السبب في تسمية طريقة إيجاد معادلته بطريقة أصغر المربعات . وقد رأينا فيما
سبق كيف استخدمت هذه الخاصية للوسط الحسابي في قياس مدى تشتت القيم
حول (الانحراف المعياري) . كذلك يمكن استخدام هذه الخاصية لخط
العلاقة المتوسطة في قياس مدى تشتت النقط حوله . والخطأ المعياري هو الذي
يقيس هذا التشتت ويرمز له بالرمز s سواء كان الخطأ المعياري الذي نحسبه
هو لمعادلة انحدار $\frac{y}{x}$ أو $\frac{x}{y}$.

ويمكن تعريف الخطأ المعياري بأنه المقياس الذي يقيس لنا درجة دقة
معادلة الانحدار في تمثيل العلاقة بين المتغيرين تمثيلاً جبرياً ، وبالطبع كلما صغر
الخطأ المعياري كلما دل ذلك على أن المعادلة تمثل العلاقة تمثيلاً صادقاً والعكس
كلما كبر الخطأ المعياري نقصت درجة دقة المعادلة في تصويرها للعلاقة .
وبمعنى آخر تكون القيم التابعة المستنبطة من معادلة الانحدار قيماً يمكن
الاعتماد عليها إذا صغر الخطأ المعياري للمعادلة ولا يمكن الاعتماد عليها إذا كبر
الخطأ المعياري .

والخطأ المعياري هو الجذر التربيعي لمتوسط مجموع مربع انحرافات القيم
النظرية للمتغير y أو المتغير x عن قيمه الواقعية . والقيم النظرية هي
المستنبطة من معادلة الانحدار أو من خط العلاقة المتوسطة بالرسم البياني .
ويمكن حساب الخطأ المعياري لمعادلة الدرجة الأولى كالآتي :

$$\text{خ}^2 \text{ لمعادلة انحدار } \frac{\text{ص}}{\text{ن}} = \frac{1}{\text{ن}} \text{ م ح ص}^2 - (\text{م ح ص} + \text{ب ح ص})$$

$$\text{خ}^2 \text{ لمعادلة انحدار } \frac{\text{ص}}{\text{ن}} = \frac{1}{\text{ن}} \text{ م ح ص}^2 - (\text{م ح ص} + \text{ب ح ص} - \text{م ح ص})$$

أما إذا كانت معادلة الانحدار من الدرجة الثانية فيمكن حساب الخطأ المعياري كالآتي :

$$\text{ح}^2 \text{ لمعادلة انحدار } = \frac{\text{ص}}{\text{ن}}$$

$$\frac{1}{\text{ن}} \text{ م ح ص}^2 - (\text{أ م ح ص}^2 + \text{ب م ح ص} + \text{ح م ح ص})$$

$$\text{ح}^2 \text{ لمعادلة انحدار } = \frac{\text{ص}}{\text{ن}}$$

$$\frac{1}{\text{ن}} \text{ م ح ص}^2 - (\text{أ م ح ص}^2 + \text{ب م ح ص} + \text{ح م ح ص})$$

في هذه المعاملات نلاحظ أن

$$\text{ن} = \text{عدد النقط} .$$

$$\text{م ح ص}^2 = \text{مجموع مربع القيم الواقعية للمتغير ص} .$$

$$\text{م ح ص}^2 = \text{مجموع مربع القيم الواقعية للمتغير ص}$$

القيم الواقعة داخل الأقواس الصغيرة هي مجموع مربع القيم النظرية لكل من ص ، س كل حسب معادلة الانحدار الخاص به .

الخطأ المعياري والارتباط :

أشرت سابقاً إلى أنه إذا وقعت جميع النقط على خط الانحدار فإن

$r^2 = 1 - \frac{C^2}{E^2}$ ، أما إذا علم لدينا الخطأ المعياري لمعادلة الانحدار $\frac{S}{C}$

فتكون $r^2 = 1 - \frac{C^2}{E^2}$. هذه العلاقة الجبرية بين معامل الارتباط S والخطأ المعياري C والانحراف المعياري E للقيم التابعة يمكن إثباتها .

والعلاقة السابقة بين معامل الارتباط والخطأ المعياري صحيحة كذلك إذا كان الانحدار من درجة أعلى من الأولى ، والفرق الوحيد هو لمجرد التمييز بين الارتباط المستقيم وغير المستقيم فنسمي مقياس الارتباط غير المستقيم دليل الارتباط ويرمز له بالرمز τ وهو يستعمل في قياس الارتباط غير المستقيم من واقع قيم غير مبوبة .

وبذلك يكون $\tau^2 = 1 - \frac{X^2}{E^2}$ في حالة انحدار $\frac{S}{C}$

، $\tau^2 = 1 - \frac{X^2}{E^2}$ في حالة انحدار $\frac{S}{C}$

ونلاحظ ان القيمة التي نحسبها لمعامل الارتباط (r) ولدليل الارتباط (τ) لا تتساوى في حالة الارتباط غير المستقيم - بمعنى انه إذا حسبنا r لقياس الارتباط بين متغيرين علاقتها من الدرجة الثانية (باستخدام قانون بيزسون مثلا) فانها تكون مقياس غير دقيق للارتباط بين هذين المتغيرين حيث ان r تحسب على أساس ان العلاقة من الدرجة الأولى والمقياس الأصح هو τ الذي يحسب من واقع المعادلات المذكورة أعلاه .

لذلك يرى بعض الاحصائيين عدم استعمال r لقياس الارتباط بين متغيرين ويقترحون استخدامه فقط في قياس درجة استقامة خطوط الانحدار

وتسميته تبعاً لذلك بمعامل الاستقامة وذلك لأنه إذا كان r كبيراً كان مقدار x^2 صغيراً مما يدل على أن الخط المستقيم يوافق جيداً النقط المعطاة ، أي أن خط الانحدار قريب من المستقيم . أما إذا كان r صغيراً كان مقدار x^2 كبيراً حول الخط المستقيم مما يدل على أن هذا الخط المستقيم لا يوافق هذه النقط جيداً ، أي أن خط الانحدار بعيد عن الاستقامة .

كما أننا إذا حسبنا r ثم حسبنا τ عند قياس العلامة بين متغيرين وكانت r أكبر من τ كان ذلك دليلاً قوياً على أن القيم المعطاة يوافقها مستقيم أفضل من أي خط آخر . والعكس إذا كانت τ أكبر من r دل ذلك على أن القيم المعطاة يوافقها منحنى أفضل من خط المستقيم ولذلك كان تشتتها حول الخط من الدرجة الثانية مثلاً الذي استخدمناه في حساب τ أقل من تشتتها حول الخط المستقيم الذي افترضناه موافقاً لتمثيل هذه القيم عند حساب المعامل r لها .

ومن ناحية أخرى إذا رسمنا شكل الانتشار وتبين لنا أن خط الاتجاه العام مستقيم يمكن حساب r ، أما إذا كان منحنى فيجب حساب τ بإيجاد المعادلة ثم حساب الخطأ المعياري لها .

على أننا نعرف أن معرفة الاتجاه العام للعلاقة بين متغيرين ليس من السهل تحديدها بدقة فهي تعتمد إلى حد كبير على الرأي الشخصي وبهذه الطريقة يدخل العامل الشخصي في أبحاثنا ويتحكم في المقياس الذي نحصل عليه لدرجة العلاقة بين المتغيرين تحت البحث .

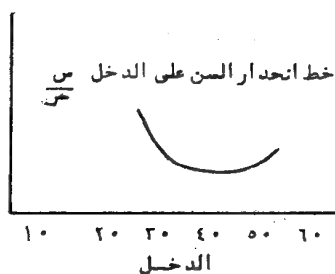
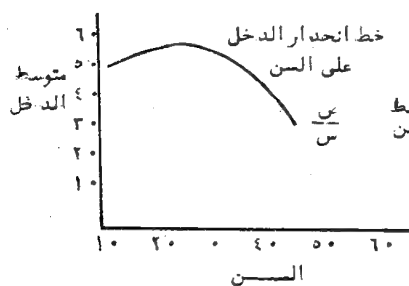
نسبة الارتباط :

عندما تكون قيم المتغيرين مبوبة في جدول تكراري مزدوج وتبين لنا أن العلاقة بينهما ليست من الدرجة الأولى فلا يمكن حساب r كمقياس دقيق للعلاقة بين المتغيرين . كذلك يكون من الصعب حساب τ لأن هذا لا بد

له من توفيق منحني يصور جبريا العلاقة بين المتغيرين (أصغر المربعان ١ .
لذلك وضع بيرسون مقياساً سماه نسبة الارتباط لاستخدامه في مثل هذه
الأحوال ويرمز له بالحرف γ .

ونسبة الارتباط بين أي متغيرين $\gamma = \frac{\text{عص}}{\text{عص}}$ حيث عص تدل على
الانحراف المعياري لتوسطات قيم ع المقابلة لمراكز فئات س ، ع هي
الانحراف المعياري لقيم ص نفسها . كذلك فإن $\gamma = \frac{\text{عص}}{\text{عص}}$ حيث عص
الانحراف المعياري لتوسطات قيم س المقابلة لمراكز فئات ص ، عص هي
الانحراف المعياري لقيم س .
مثال ٦٤ :

السن الدخل	صفر -	- ١٠ -	- ٢٠ -	- ٣٠ -	- ٤٠ -	لمجموع	متوسطات السن
- ٣٠ -	٤				٨	١٢	٣١٦٧
- ٣٥ -	٨				٤	١٢	١٨٣٣
- ٤٠ -	١٤	١٢		٨		٤٤	١٧٧٣
- ٤٥ -		٦	١٤			٢٠	٢٢٠٠
- ٥٠ -			١٢			١٢	٢٥٠٠
المجموع	٢٦	١٨	٣٦	٨٠	١٢	١٠٠	
متوسطات الدخل	٣٩٤٢	٤٤١٧	٤٧٧٨	٤٢٥٠	٣٤١٧		



يتضح من هذا الرسم ان العلاقة بين هذين المتغيرين ليست من الدرجة الاولى وحيث أن القيم الخاصة بها مبوبة فتكون نسبة الارتباطى دي أحسن مقياس لدرجة العلاقة بينها . ولحسابى تتبع الخطوات الآتية :

اولا - حساب ع

متوسطات الدخل	ك	متوسطات الدخل في ك	ع من الوسط الحسابي	ع ك	ك ² ع ك
٣٩٤٢	٢٦	١٠٢٤ر٩٢	٣٤٨-	٩٠٤٨-	٣١٤٨٧
٤٤١٧	١٨	٧١٥٠٦	١٢٧+	٢٢٨٦+	٢٩٠٣
٤٧٧٨	٣٦	١٧٢٠٠٨	٤٨٨+	١٧٥٦٨+	٥٨٧٣٢
٤٢٥٠	٨	٣٤٠٠٠	٤٠-	٣٢٠-	١٢٨
٣٤١٧	٦٠٢	٤١٠٠٤	٨٧٣-	١٠٤٧١-	٩١٤٥٥٠
المجموع	١٠٠	٤٢١٠١			١٨٤٧٠٥

$$\text{ع (الوسط الحسابي)} = \frac{٤٢١٠١}{١٠٠} = ٤٢١.٠١$$

$$\text{ع} = \sqrt{\frac{١٨٤٧٠٥}{١٠٠}} = \sqrt{١٨٤٧.٠٥} = ٤٢٩.٨٠$$

ثانيا - حساب ع من

الغلات	ك	م	ك ² م	ع من الوسط الحسابي	ع ك	ك ² ع ك
٣٠-	١٢	٣٢٥	٣١٠	١٠٤-	١٢٤٨-	١٢٩٧٩١
٣٥-	١٢	٣٧٥	٤٥٠	٥٤-	٦٤٨-	٣٤٩٩٢
٤٠-	٤٤	٤٢٥	١٨٧٠	٤-	١٧٦-	٧٠٤
٤٥-	٢٠	٤٧٥	٦٥٠	٦٠+	٩٢٠+	٤٢٣٢٠
٥٠-	١٢	٥٢٥	٦٣٠	٦٠+	١١٥٢+	١١٠٥٦٢
المجموع	١٠٠		٤٢٩٠			٣١٨٤

$$\text{ع من} = \frac{٤٢٩٠}{١٠٠} = ٤٢.٩$$

$$٥٦٤٣ = \frac{٣١٨٤}{١٠٠} \vee = \frac{٣١٨٤}{١٠٠} \vee = \text{ع س}$$

$$\text{نسبة الارتباط ي} = \frac{٤٢٩٨}{٥٦٤٣} = ٧٦٢ \text{ ر.}$$

ومما يجب التنويه به ان نسبة الارتباط مقياس أعم لحساب الارتباط بين ظاهرتين سواء كانت العلامة بينها مستقيمة أو غير مستقيمة ، بخلاف معامل الارتباط الذي لا يجب أن نستعمله الا إذا تأكدنا أن الانحدار بين المتغيرين مستقيم أو قريب جداً من الاستقامة .

استنباط قوانين خطوط الانحدار من الدرجة الاولى :

عندما يكون الانحدار مستقيماً تكون المعادلة التي تصور العلاقة بين الظاهرتين تصويراً جبرياً هي :

$$(١) \quad \text{ص} = \text{م} + \text{س} + \text{ب}$$

$$(٢) \quad \text{أو} \quad \text{س} = \text{م} - \text{ص} + \text{ب} -$$

ولايجاد المعادلة الأولى نتبع طريقة المربعات الصغرى حيث نستخدم المعادلتين الآتيتين لايحاد قيمة كل من م ، ب :

$$(٣) \quad \text{مح} \quad \text{ص} = \text{م مح} + \text{ن ب}$$

$$(٤) \quad \text{مح س} = \text{ص} = \text{م مح} + \text{ب مح} + \text{س}$$

ونلاحظ أنه يمكن وضع المعادلة (٣) في الشكل التالي :

$$\text{ن} \quad \text{ص} - = \text{ن م} - \text{س} + \text{ب}$$

$$\text{بالقسمة على ن ينتج ان} \quad \text{ص} - = \text{م} - \text{س} + \text{ب}$$

$$(٥) \quad \therefore \text{ب} = \text{ص} - - \text{م س} -$$

كذلك يمكن التعويض في المعادلة (٤) عن قيمة $مح س ص$ من القانون الثاني للارتباط (بيرسون) ، وعن $س٢$ من قانون الانحراف المعياري .

$$ر = \frac{مح س ص - ن س - ص}{\begin{matrix} ن ع & ع ن \\ س & ص \end{matrix}}$$

$$\therefore مح س ص = ن ر ع + ن س - ص$$

$$ع٢ = س - \frac{مح س٢}{ن}$$

$$\therefore ع٢ = س - \frac{مح س٢}{ن}$$

$$\therefore مح س٢ = ن ع٢ + ن س - ص$$

.. بالتعويض في المعادلة (٤) نحصل على :

$$ن ر ع س ع ص + ن س - ص = ن م ع٢ س + ن م س - ص + (ن س - ص - م س)$$

$$\therefore ن ر ع س ع ص + ن س - ص = ن م ع٢ س + ن م س - ص + ن س - ص - م س$$

$$\therefore ن ر ع س ع ص = ن م ع٢ س$$

$$\therefore م = \frac{ن ر ع س ع ص}{ن ع٢ س}$$

(٦)

$$\frac{ع}{ع} ر = م . .$$

بالتعويض بقيمة ب وقيمة م في معادلة الانحدار $ص = م س + ب$

$$\therefore ص = ر \frac{ع}{ع} س + ص - ر \frac{ع}{ع} س$$

$$\therefore ص - ص = -ص - ر \frac{ع}{ع} (س - س) \quad (٧) \text{ وهي المعادلة}$$

التي نستعملها لإيجاد انحدار $\frac{ص}{س}$ عندما يكون معلوماً لدينا المقاييس س- ،
ص- ، ع ص ، ع س ، ر خاصة عندما تكون القيم مبوبة في توزيع تكراري
مزدوج

$$\text{وبالمثل يمكن استنباط معادلة انحدار } \frac{س}{ص} \text{ وهي } س - س = -ص$$

$$ر \frac{ع}{ع} (ص - ص)$$

قانون الخطأ المعياري لمعادلة الانحدار من الدرجة الاولى :

نفترض أن لدينا ظاهرتان س ، ص وان قيمها المتناظرة هي كالآتي :

ص ۱

۱۵

ص ۲

۲۳

ص ۳

مس ۴

من

مس و

...

..

ص ن

مس

واننا وجدنا معادلة المنحدر $\frac{ص}{س}$ فكانت $ص = م س + ب$ وبذلك إذا

عوضنا عن س بقيهما نحصل على القيم النظرية للمتغير ص التي تقابل قيمه الواقعية ، وتكون هذه القيم النظرية هي م س١ + ب م س٢ + ب م س٣ + ب .

وبذلك يكون الفرق بين كل قيمة واقعية للمتغير x وقيمتها النظرية المقابلة هي :

ص_۱ - (م س_۱ + ب) - ص_۲ - (م س_۲ + ب) - ص_۳ -
 (م س_۳ + ب) - ص_۴ - (م س_۴ + ب) - ص_ن - (س_ن + ب)
 وبتربيع هذه الفروق نحصل على :-

[illegible]

وهكذا حتى نصل الى الفرق النوني وبالجمع نحصل على :

$$\text{مجد ص}^2 - \text{م}^2 \text{م ح س ص} + \text{م}^2 \text{م ح س}^2 - \text{ب}^2 \text{م ح ص} + \text{م}^2 \text{م ب} \\ \text{م ح س} + \text{ن ب}^2.$$

وبالتعويض عن مد ص بما تساويه م مد س^٢ + ب مد س
وعن مد ص بما تساويه م مد س + ن ب نحصل على :

$$\text{مد ص}^2 - \text{م مد س ص} - \text{ب مد ص} - \text{م مد س}^2 - \text{م ب مد س} - \text{ن ب}^2 + \text{م مد س}^2 + \text{م ب مد س} - \text{م ب مد س}.$$

وهذا يساوي بعد الاختصار مد ص^٢ - م مد س ص - ب مد ص
أي مد ص^٢ - (م مد س ص + ب مد ص)

وحيث ان مربع خطأ المعادلة يساوي متوسط هذه الفروق :

$$\text{مربع خطأ المعادلة} = \frac{1}{n} (\text{مد ص}^2 - (\text{م مد س ص} + \text{ب مد ص}))$$

$$\text{وبتحليل القيمة} \quad \frac{\text{مد ص}^2 - (\text{م مد س ص} + \text{ب مد ص})}{n} \quad \text{أي}$$

وبالتعويض عن مد ص^٢ بما تساويه من قانون الانحراف المعياري ع ص =

$$\frac{\text{مد ص}^2}{n} - \text{ص}^2 \quad \text{وعن م مد س ص بما تساويه من قانون الارتباط ر} =$$

$$\frac{\text{مد س ص} \quad \text{ن س} - \text{ص} - \text{ن ع س ع ص}}{n}$$

وعن مد ص بما تساويه ن ص - وعن ب بما تساويه ص - م س - .

$$= \text{ح}^2.$$

$$\frac{\text{ن ع}^2 \text{ ص} + \text{ن ص}^2 - \text{ن م ر ع س ع ص} - \text{ن م س}^2 \text{ ص} - \text{ن ص}^2 - \text{ن م س}^2 \text{ ص} - \text{ن م س}^2 \text{ ص}}{n}$$

$$\therefore \text{ع}^2 \text{ ص} - \text{م ر ع س ع ص} = \text{مربع خطأ المعادلة}$$

وبالقسمة على ع^٢ ص نحصل على :

$$١ - م ر = \frac{ع س}{ع ص} = \frac{\text{مربع خطأ المعادلة}}{ع^٢ ص}$$

$$\text{ولكن م ر} = \frac{ع ص}{ع س}$$

$$\therefore ١ - م ر = \frac{ع ص}{ع س} \times \frac{ع س}{ع ص} = \frac{\text{مربع خطأ المعادلة}}{ع^٢ ص}$$

$$\therefore ١ - م ر^٢ = \frac{\text{مربع خطأ المعادلة}}{ع^٢ ص}$$

$$\therefore م ر^٢ = ١ - \frac{\text{مربع خطأ المعادلة}}{ع^٢ ص}$$

$$\text{أي م ر}^٢ = ١ - \frac{ع خ^٢}{ع^٢ ص}$$

دلالة معامل الارتباط :

ان اعتمادنا على معامل الارتباط لقياس درجة الارتباط بين ظاهرتين يمكن أن يضلنا حيث قد يحدث تلازم بين قيم الظاهرتين لمجرد الصدفة . ولا شك أن تأثير عامل الصدفة على النتيجة التي نصل اليها عند قياس الارتباط يزداد كثيراً كلما نقصت عدد الحالات التي أجرينا عليها البحث ، وبذلك تكون دلالة معامل الارتباط دلالة ضعيفة . فلو فرضنا اننا درسنا حالتين فقط (وحدتين فقط) فان الارتباط بينهما لا بد أن يكون كاملاً مهما كانت قيم المتغيرين ومن الواضح أن هذه النتيجة قد تكون بعيدة كلياً عن الحقيقة - إلا انه كلما زاد عدد أزواج القيم المتناظرة التي حسب الارتباط على أساسها

فان احتمال الحصول على معامل ارتباط قوي بين الظاهرتين بطريق الصدفة يكون احتمال ضعيف ويزداد ضعف الاحتمال كلما زاد عدد أزواج قيم الظاهرتين - فاذا توصلنا إلى معامل ارتباط ٠.٤ من دراسة ١٠٠٠ حالة تكون دلالة هذا المعامل أكبر بكثير من دلالة معامل ارتباط ٠.٧ نتج عن دراسة عشر حالات فقط .

ولقد أعد الجدول الآتي ليبين احتمالات الحصول على قيم مختلفة لمعامل الارتباط بطريق الصدفة حسب عدد أزواج القيم المتناظرة :

احتمال الحصول على قيمة r المبينة في هذا الجدول بطريق الصدفة				عدد أزواج القيم المتناظرة
٠.١	٠.٢	٠.٥	٠.٨	
٠.٧٦- r	$r = ٠.٧٢$	$r = ٠.٦٣$	$r = ٠.٥٥$	١٠
٠.٦٤- r	٠.٥٩- r	٠.٥١- r	٠.٤٤- r	١٥
٠.٥٦- r	٠.٥٢- r	٠.٤٤- r	٠.٣٨- r	٢٠
٠.٤٩- r	٠.٤٥- r	٠.٣٨- r	٠.٣٢- r	٢٧
٠.٤٥- r	٠.٤١- r	٠.٣٥- r	٠.٣٠- r	٣٢
٠.٤٢- r	٠.٣٨- r	٠.٣٢- r	٠.٢٨- r	٣٧
٠.٣٩- r	٠.٣٦- r	٠.٣٠- r	٠.٢٦- r	٤٢
٠.٣٥- r	٠.٣٣- r	٠.٢٨- r	٠.٢٣- r	٥٠
٠.٣٣- r	٠.٣٠- r	٠.٢٥- r	٠.٢١- r	٦٠
٠.٣٠- r	٠.٢٨- r	٠.٢٣- r	٠.٢٠- r	٧٠
٠.٢٨- r	٠.٢٦- r	٠.٢٢- r	٠.١٨- r	٨٠
٠.٢٧- r	٠.٢٤- r	٠.٢٠- r	٠.١٧- r	٩٠
٠.٢٦- r	٠.٢٣- r	٠.٢٠- r	٠.١٦- r	١٠٠

ويمكن اختيار دلالة معامل الارتباط بطريقة أخرى ذلك أننا غالباً في قياس الارتباط نعتمد على عينة من المجتمع موضوع البحث ، وبذلك على فرض

عدم وجود ارتباط بين الظاهرتين موضوع البحث فان معامل الارتباط لو حسب للمجتمع كله وليس للعينة فقط كانت النتيجة تساوي صفراً . فاذا اخذنا عدداً كبيراً جداً من العينات من هذا المجتمع وحسبنا لكل منها معامل الارتباط ثم وزعنا هذه المعاملات في توزيع تكراري لحصلنا على توزيع معتدل متوسطه أي ان معامل الارتباط المتوسط (الواقعي للمجتمع) = صفر . ومن الواضح ان الانحراف المعياري لهذا التوزيع يدل على حجم الاخطاء (الفروق) التي يحتمل ان تظهر بين نتائج العينات والنتيجة الواقعية (صفر) ، وبعبارة اخرى يكون هذا الانحراف المعياري هو الخطأ المعياري لمعامل الارتباط .

$$\frac{1}{\sqrt{1 - n}} = \text{والخطأ المعياري لمعامل الارتباط}$$

وتبعاً لخصائص التوزيع المعتدل يكون الفرق الذي يساوي ضعف هذا الخطأ في الحد الأقصى فرق راجع الى الصدفة (بدرجة ثقة ٩٥٤ ٪) ، أما الفرق الذي يساوي أكثر من ضعف الخطأ المعياري يكون فرقاً جوهرياً ، وبذلك اذا كان معامل الارتباط الذي نحصل عليه من العينة (يؤخذ المعامل نفسه لأن فرقه عن الصفر = ر) أقل من ضعف الخطأ المعياري نستطيع أن نستنتج ان هناك احتمال كبير (٩٥٤ ٪) في أن يكون هذا المعامل راجعاً الى الصدفة ، واذا كان معامل الارتباط اكبر من ضعف الخطأ المعياري نستطيع أن نستنتج ان هناك احتمال صغير (١٠٠ - ٩٥٤ = ٤٦ ٪) في أن يكون هذا المعامل راجعاً الى الصدفة بمعنى أن هذا المعامل له دلالة ومعنويته في قياس الارتباط بين الظاهرتين :

مثال ٦٥ :

فإذا حصلنا على معامل ارتباط ٧ ر من دراسة ١٠٠ حالة

$$\frac{1}{\sqrt{99}} = \frac{1}{\sqrt{1 - n}} = \text{معامل الارتباط}$$

$$= 100.5 \text{ ر}$$

$$\therefore 2 \text{ ع} = 2 \times 100.5 = 201.0.$$

وحيث ان معامل الارتباط ٧ ر أكبر بكثير من هذا الرقم ، فلا يمكن بأي حال من الأحوال ان يكون هذا المعامل راجعاً الى الصدفة ، أي ان هذا المعامل يدل دلالة حقيقية على وجود الارتباط بين الظاهرتين موضوع الدرس .

تمارين

(١) من القيم الآتية للمتغيرين س ، ص أوجد معادلتى الانحدار بطريقتين :

س ٩٢ ١٠٦ ١٠٣ ١٣٨ ١٧٥ ٢٠١ ١٦٩ ١٤٦ ١٦ ١٧
ص ١٠٥ ٩٢ ٩٧ ١١٥ ١٢٥ ١٣ ١٤٣ ١٥ ١٣٩ ١٧٤

$$(\text{ص} = ٤٦٥.٠ + \text{س} + ٥٩)$$

$$(\text{س} = ٩١٤.٠ + \text{ص} + ٣)$$

(٢) احسب معامل الارتباط بين س ، ص بطريقة سيرمان :

س ١٠ ١٢ ١٣ ١٦ ١٧ ٢٠ ٢٦ ١١
ص ١٩ ٢٢ ٢٣ ٢٤ ٢٢ ٢٧ ٢٢ ٣٣

(٣) من الجدول التالي اوجد معامل الارتباط بين الظاهرتين س ، ص :

م	س				المجموع
	٢ -	٤ -	٦ -	٨ -	
١٠ -	٢	٤	-	-	٦
٢٠ -	١	٦	٢	-	٩
٣٠ -	١	٤	٤	-	٩
٤٠ -	-	٣	٨	٣	١٤
٥٠ - ٦٠	-	-	٥	٧	١٢
المجموع	٤	١٧	١٩	١٠	٥٠

ثم اوجد معادلي الانحدار - اختبر دلالة معامل الارتباط .

٤ - الجدول الآتي يبين توزيع ١٠٠ أسرة تبعاً لعدد حجرات السكن (ص) والدخل الشهري (س) .

س	ص					المجموع
	١	٢	٣	٤	٥	
٥ -	٥	٥	٣	١	-	١٤
١٠ -	٣	٧	١٠	٤	-	٢٤
١٥ -	٢	٩	٢٢	٥	-	٣٨
٢٠ -	-	٤	٨	٣	١	١٦
٢٥ - ٣٠	-	٢	٢	-	٤	٦
المجموع	١٠	٢٧	٤٥	١٣	٥	١٠٠

المطلوب :

- ١ - رسم شكل انتشار يبين انحدار الايجار على عدد الغرف والتعليق عليه
- ٢ - حساب معامل الارتباط واختبار دلالة .
- ٣ - ايجاد معادلة انحدار الايجار على عدد الغرف .
- ٤ - حساب خطأ المعادلة - فسر النتيجة .
- ٥ - تقدير متوسط الدخل الذي يدفع لأربع غرف وتحقيق النتيجة من الجدول .

الفصل الثالث عشر

ع

الارقام القياسية

نشأ الاهتمام بموضوع الارقام القياسية نتيجة الحاجة إلى مقياس محدد يقيس التغيرات في قيمة النقود. على ان ذلك ليس هو المجال الوحيد لاستخدام الارقام القياسية ، إذ قد اتسع مجال استخدامها حتى أصبح يشمل كثيراً من الظواهر الاقتصادية وغير الاقتصادية .

والرقم القياسي هو مقياس احصائي لقياس التغيرات التي تحدث في متغير واحد أو في مجموعة من المتغيرات خلال الزمن أو من منطقة جغرافية إلى أخرى أو على اساس اختلاف الدخول والمهن ... الخ .

وباستعمال الارقام القياسية نستطيع مثلاً مقارنة تكلفة الطعام أو أي بند آخر من بنود الاتفاق هذا العام بالنسبة للعام السابق ، أو نستطيع مقارنة انتاج الصلب في عام معين في منطقة معينة من الدولة بالنسبة لمنطقة أخرى في نفس العام . وبالرغم من ان الارقام القياسية تستخدم اساساً في قياس التغيرات الخاصة بالظواهر الاقتصادية ، الا انه يمكن تطبيقها كذلك في نواحي أخرى كثيرة ، مثلاً في مجال التربية نستطيع باستخدام الارقام القياسية في مقارنة ذكاء الطلاب في مناطق الدولة او في سنوات مختلفة .

وتهتم كثير من الادارات الحكومية والخاصة بتركيب أرقام قياسية يمكن استعمالها في توقع المستقبل بالنسبة للاموال التجارية والنشاطات

الاقتصادية المختلفة ، ولذلك نجد في كثير من الدول أرقاماً قياسية للاجور
واخرى للانتاج بأنواعه المختلفة واخرى للبطالة ... الخ . ولعل أكثر هذه
الارقام شيوعاً هو الرقم القياسي لنفقة المعيشة أو كما يسمى أحياناً الرقم
القياسي لاسعار السلع الاستهلاكية ، وفي بعض الدول تحتوي عقود العمل
الجماعية على نص بزيادة الاجور تلقائياً كلما دلت هذه الارقام القياسية على
الارتفاع .

ويمكن أن نعطي هنا أمثلة لبعض التطبيقات العملية للارقام القياسية :

١ - يقاس التغير في الأسعار في فترة زمنية معينة وذلك لاكتشاف سبب
التغير وأثره على النشاط الاقتصادي حتى يمكن التحكم فيه .

٢ - قياس التغير في نفقة المعيشة وخاصة بالنسبة للطبقة العاملة حتى
يمكن البت في مطالبهم الخاصة بالاجور .

٣ - قياس التغير في مستوى معيشة طبقة معينة من الناس ، وذلك
بقسمة الرقم القياسي لدخولهم النقدي على الرقم القياسي لنفقة
معيشتهم فنحصل على رقم قياسي للأجر الحقيقي يبين لنا التغير في
مستوى معيشة هذه الطبقة فنستطيع أن نقرر مساعدتهم إذا تبين
لنا أن التغير في حالتهم يدعو إلى ذلك .

٤ - قياس التغير في حجم الانتاج الصناعي والزراعي والموجودات في
المخازن والتجارة الخارجية والمبيعات .

٥ - قياس التغير في عدد المشتغلين عامة وفي كل ناحية من نواحي النشاط
الاقتصادي على حدة ، وكذلك التغير في عدد المتعطلين .

تكوين الرقم القياسي

يواجه الاحصائي المشاكل الآتية عند تكوين رقم قياسي :

- ١ - اختيار المتغيرات التي يشملها الرقم .
- ٢ - اختيار المصادر التي يجمع منها البيانات اللازمة .
- ٣ - اختيار الفترة أو المكان الذي يعتبر أساساً لقياس التغير
- ٤ - اختيار الصيغة التي تستخدم في حساب الرقم القياسي حيث توجد عدة صيغ سيأتي شرحها فيما بعد .
- ٥ - اختيار الأوزان التي تستخدم في الترجيح إذا اتفقنا على صيغة يكون فيها الرقم القياسي مرجحاً .
- ٦ - تنظيم العمل بحيث يمكن حساب الرقم سريعاً حتى تتحقق فائدته العملية إذ لا فائدة عملية لرقم قياسي ينشر بعد فوات الوقت المتعلق به .

ولشرح هذه المشاكل سنقتصر في دراستنا على الأرقام القياسية للأسعار مع ملاحظة أن ما ينطبق على هذا النوع من الأرقام القياسية ينطبق أيضاً على الأنواع الأخرى مع بعض التعديلات الطفيفة التي تناسب كل نوع .

المشكلة الأولى :

عند تكوين رقم قياس للأسعار يجب أن نحدد أنواع السلع التي يشملها الرقم ، وبالطبع يتوقف هذا التحديد على نوع الرقم ؛ فإذا كنا نبغي رقماً يقيس التغير في نفقة المعيشة لطبقة معينة من الناس يجب أن نقتصر على السلع التي تستهلكها هذه الطبقة ، وإذا كنا نبغي رقماً يقيس التغير في نفقات الإنتاج الصناعي اقتصرنا على السلع التي تعتبر مواداً أولية في الصناعة . وبذلك يكون أول ما يجب أن نراعيه عند تكوين رقماً قياسياً أن نختار عينة ممثلة تمثيلاً صادقاً للموضوع الذي نريد دراسته ، ولكي تكون العينة

ممثلة يجب أن تشمل عدداً لا بأس به من السلع ، وليس هناك قاعدة محدودة يمكن أن نراعيها عند تحديد عدد السلع وان كان يجب أن نلاحظ أن لا يزيد العدد كثيراً لأن في ذلك تعقيد للعمليات الحسابية ، الامر الذي يؤخر ظهور النتائج ، والا يقل العدد كثيراً حيث تصبح العينة نتيجة لذلك غير ممثلة تمثيلاً صادقاً .

ويحسن تقسيم السلع التي تدخل في الرقم إلى مجموعات ، تتميز المفردات في كل منها بصفات خاصة ذات أهمية في الناحية التي يتناولها الرقم القياسي الذي نريد تكوينه . فمثلاً عند تكوين رقم قياسي لنفقة المعيشة نقسم السلع إلى مجموعة اللحوم ، والخضروات ، والفواكه ، والملابس .. الخ ولا يجب أن نأخذ جميع السلع التي تتكون منها كل مجموعة فيكفي أن نأخذ بعضها والتي تكون شائعة الاستعمال لدى الطبقة التي يراد قياس نفقة معيشتها والتي تمثل الاتجاهات المختلفة في الاسعار بمعنى انه إذا كان هناك عدة سلع تسير اسعارها سوية في اتجاه واحد سواء بالارتفاع أو الانخفاض يكفي أن نأخذ منها واحدة أو اثنتين ، أما السلع التي تتغير اسعارها في اتجاهات مختلفة فيحسن أن نأخذها كلها .

المشكلة الثانية :

عند اختيار المصادر التي تجمع منها البيانات الخاصة بالرقم القياسي يجب أن نلجأ إلى التجار الذين نعتقد ان افراد الطبقة التي نريد قياس نفقة معيشتها يشتركون منهم حاجاتهم . كذلك يجب أن نتأكد من أن الاسعار التي نجعلها هي عن السلع التي قررنا إدخالها في الرقم مع ملاحظة الانواع المختلفة من السلعة الواحدة ، فإذا أردنا أن نجمع بيانات عن أسعار السيارات فنلاحظ أنواعها المختلفة وكذلك ان النوع الواحد يختلف سعره حسب السنة التي أنتج فيها ، فسيارة موديل ١٩٦٥ يختلف سعرها عن سيارة موديل ١٩٤٠ . كذلك

يجب أن نراعي تغير استعمال السلع من وقت الى آخر ، فسلعة ما قد تكون شائعة الاستعمال في وقت معين ولكن يبطل استعمالها في وقت آخر ، الأمر الذي يجعلنا نعمل دائماً على تعديل الرقم القياسي من وقت الى آخر (جميع الدول عدلت الرقم القياسي لنفقة المعيشة فيها حتى يتفق مع تغير طرق المعيشة) .

كذلك عند جمع البيانات عن الاسعار يجب ملاحظة الآنية في ذلك بمعنى أن نجمع الاسعار في نفس الوقت ولكي ندرك أهمية ذلك نتذكر التغيرات التي تصيب الاسعار من وقت إلى آخر وعند تحديد الوقت يجب مراعاة الفترة التي يقبل الأفراد فيها فعلاً على الشراء .

ومراعاة الدقة في جمع البيانات عن الأسعار من الأهمية بمكان كبير إذ إن أي خطأ لا بد أن يظهر تأثيره في النتيجة النهائية للرقم القياسي وذلك بعكس الخطأ في الأوزان التي ترجح بها فإنه لا يؤثر كثيراً في النتيجة، وذلك لأن الوزن المضروب في السعر أو في منسوبه يوجد في البسط وفي المقام أيضاً . أما إذا أخطأنا في السعر فإن البسط يتغير وحده دون المقام فيكون التغير في قيمة الكسر أكثر منه في الحالة الأولى . وعلى ذلك يجب أن نغني عناية تامة عند جمع بيانات الاسعار .

المشكلة الثالثة :

لتركيب الرقم القياسي لأية ظاهرة تكون نسبة مئوية بين القيمة المقارنة لهذه الظاهرة أو قيمتها في الفترة التي نريد مقارنتها والقيمة الأخرى المعتبرة أساساً للمقارنة ، فالرقم القياسي لسعر القمح في سنة ١٩٦٥ مثلاً هو عبارة عن النسبة المئوية لسعر القمح في سنة ١٩٦٥ الى سعره في سنة أخرى تتفق على اعتبارها أساساً للمقارنة .

وبذلك عند إنشاء أي رقم قياسي يجب أن نحدد الأساس الذي سنتخذ

لترتيب الرقم ، فإذا كانت المقارنة على أساس الزمن يجب أن نحدد الفترة التي نعتبرها أساساً ، وإذا كانت المقارنة على أساس المكان يجب أن نحدد المكان الذي نعتبره أساساً .

وعند اختيار السنة الأساس نراعي ان تكون السنة تتميز بالاستقرار في الاسعار بمعنى أن لا تكون سنة شاذة من ناحية النشاط الاقتصادي فلا تكون سنة كساد أو سنة رواج . ولعل هذا هو السبب في أن معظم الدول كانت قبل الحرب العالمية الثانية تستخدم سنة ١٩١٣ أساساً للأرقام القياسية التي تكونها . ثم أصبحت تستخدم سنة ١٩٣٨ أساساً بعد الحرب العالمية الثانية حيث ان عام ١٩١٣ أصبح بعيداً جداً والمقارنة معه تكون مضللة .

ولكن بالرغم من أن سنة ما قد تكون مناسبة كأساس في مرحلة تاريخية معينة فإنها تصبح غير مناسبة بمرور الزمن ولهذا يحسن تغيير فترة الأساس إلى فترة قريبة من الفترة التي نقارنها . ويرجع ذلك الى عدة أسباب منها تغيير طرق الاستهلاك وتغير أنواع السلع .

ويتبع أحياناً طريقة السلسلة في تكوين الرقم القياسي بمعنى أن يحسب الرقم على أساس متحرك فلا تكون هناك فترة ثابتة تقارن الاسعار على أساسها ، وإنما تكون فترة الأساس متغيرة حيث تكون سلسلة من الأرقام القياسية كل سنة بالنسبة الى سابقتها وذلك إذا لم نضمن ثبات الظروف الاقتصادية ، وبذلك يصبح الرقم القياسي مرناً مرونة لا تتوفر في الرقم القياسي للقاعدة الثابتة حيث لا يتمشى في هذه الحالة مع الظروف الاقتصادية التي تحيط بالسلع الداخلة في تركيب الرقم .

المشكلة الرابعة :

إذا قسمنا سعر أي سلعة في سنة ما على سعرها في سنة الأساس و ضربنا الناتج في ١٠٠ فإننا نحصل على ما يسمى بنسوب السعر ، وهو اصطلاح نستخدمه في حالة تكوين رقم قياسي لسلعة واحدة فقط . أما إذا أدخلنا في تكوين

الرقم أكثر من سلعة واحدة فإننا نسمي الرقم الناتج بالرقم القياسي
ولتكوين رقم قياسي هناك طريقتان: طريقة التجميع وطريقة المناسيب.
وهناك صيغ مختلفة لكل من الطريقتين

طريقة التجميع - الرقم التجميعي البسيط للأسعار :

لتركيب هذا الرقم يقسم مجموع أسعار السلع في سنة المقارنة على مجموع
أسعارها في سنة الأساس ونضرب الناتج في ١٠٠ :

$$\text{أي } \frac{\text{مجموع} ١٠٠ \times}{\text{مجموع}} - ١ -$$

مثال ٦٦ :

السلع	أ	ب	ج	د
الأسعار في سنة ١٩٣٠	٢٠	١٢	٢٢	٩
الأسعار في سنة ١٩٥٠	٤١	١٩	٢٩	١١

$$\text{الرقم التجميعي البسيط} = \frac{١١ + ٢٩ + ١٩ + ٤١}{٩ + ٢٢ + ١٢ + ٢٠} \times ١٠٠$$

وميزة هذا الرقم هي البساطة في حسابه ولكن هذه الميزة هي عيب
يؤخذ عليه في نفس الوقت ذلك لأن في استعمال هذه الصيغة نعتبر جميع السلع
متعادلة في الأهمية من حيث تأثيرها على المتوسط العام للأسعار وهذا اعتبار
خاطئ فلكل سلعة أهميتها وتأثيرها الذي يتناسب مع تلك الأهمية .

الارقام التجميعية المرجحة

لتصحيح الرقم التجميعي البسيط ترجح أسعار السلع المختلفة بأوزان

تتناسب وأهمية هذه السلع فتستخدم الكميات المنتجة أو المستهلكة من هذه السلع كأوزان ، ويصح أن نستعمل كميات سنة المقارنة أو سنة الأساس أو متوسطها . وهناك أنواع مختلفة من الأوزان سيأتي شرحها فيما بعد والصيغ المختلفة للأرقام المرجحة هي :

١ - الرقم التجمعي المرجح بكميات سنة الأساس $\frac{\text{م.ع. ك.}}{\text{م.ع. ك.}} \times 100$ (رقم لاسير)

٢ - الرقم التجمعي المرجح بكميات سنة المقارنة $\frac{\text{م.ع. ك.}}{\text{م.ع. ك.}} \times 100$ (رقم باشه)

٣ - الرقم التجمعي المرجح بالوسط الحسابي لكميات الأساس والمقارنة . $\frac{\text{م.ع. (ك. + ك.)}}{\text{م.ع. (ك. + ك.)}} \times 100$ (الرقم القياسي لمارشال وادجورث)

٤ - الرقم التجمعي المرجح بالوسط الهندسي لكميات الأساس والمقارنة . $\frac{\sqrt{\text{م.ع. ك.} \times \text{م.ع. ك.}}}{\sqrt{\text{م.ع. ك.} \times \text{م.ع. ك.}}} \times 100$

٥ - الرقم الأمثل وهو عبارة عن الوسط الهندسي للرقمين الأول والثاني . (رقم فيشر)

$$\sqrt{\frac{\text{م.ع. ك.}}{\text{م.ع. ك.}} \times \frac{\text{م.ع. ك.}}{\text{م.ع. ك.}}} \times 100$$

١٠٠ × ١٤ / ١٠٠ = ١٤

مثال ٦٧ :

السلعة	كمية سنة الأساس	كمية سنة المقارنة	أسعار سنة الأساس	أسعار سنة المقارنة
أ	٢٥٥	٢٨٠	٢٠	٤١
ب	١٨٥	٢٢٥	٢٢	١٩
ج	٣٠٥	٤١٨	٢٢	٢٩
د	١١٠	١٢٤	٩	١١

الرقم ١ =

$$100 \times \frac{110 \times 11 + 305 \times 29 + 185 \times 19 + 255 \times 41}{110 \times 9 + 305 \times 22 + 185 \times 12 + 255 \times 20}$$

الرقم ٢ =

$$100 \times \frac{125 \times 11 + 418 \times 29 + 235 \times 19 + 280 \times 41}{125 \times 9 + 418 \times 22 + 235 \times 12 + 280 \times 20}$$

الرقم ٣ =

$$100 \times \frac{(125+110)11 + (418+305)29 + (235+185)19 + (280+255)41}{(125+110)9 + (418+305)22 + (235+185)12 + (280+255)20}$$

الرقم ٤ =

$$100 \times \frac{125 \times 11 + 11 + 418 \times 305 \times 29 + 235 \times 185 \times 19 + 280 \times 255 \times 41}{120 \times 11 + 9 + 418 \times 305 \times 22 + 235 \times 185 \times 12 + 280 \times 255 \times 20}$$

الرقم / ٥ يمكن حسابه بإيجاد الوسط الهندسي لنتيجة الرقم الاول والثاني .

طريقة المناسيب :

يمكن حساب الرقم القياسي على أساس الوسط لمناسيب اسعار السلعة الداخلة في تكوين الرقم ، والوسط اما أن يكون حسابياً او هندسياً أو توافقياً ، وهذه المتوسطات اما ان تكون بسيطة او مرجحة .

ولحساب الرقم القياسي تبعاً لهذه الطريقة نبدأ بحساب منسوب السعر لكل سلعة تدخل في تكوين الرقم .

$$\text{منسوب سعر السلعة أ} = \frac{١ع}{.ع} \times ١٠٠ = ١س$$

$$\text{منسوب سعر السلعة ب} = \frac{١ع}{.ع} \times ١٠٠ = ٢س$$

$$\text{منسوب سعر السلعة ج} = \frac{١ع}{.ع} \times ١٠٠ = ٣س$$

$$\text{منسوب سعر السلعة د} = \frac{١ع}{.ع} \times ١٠٠ = ٤س$$

وبذلك يكون الرقم القياسي لأسعار هذه السلع باستخدام الوسط الحسابي البسيط وهو :

$$\frac{١س + ٢س + ٣س + ٤س}{٤}$$

والرقم القياسي لأسعار هذه السلع باستخدام الوسط الهندسي البسيط هو :

$$\sqrt[4]{س_١ \times س_٢ \times س_٣ \times س_٤}$$

والرقم القياسي لأسعار هذه السلع باستخدام الوسط التوافقي البسيط هو :

$$\frac{4}{\frac{1}{س_١} + \frac{1}{س_٢} + \frac{1}{س_٣} + \frac{1}{س_٤}}$$

هذه المتوسطات البسيطة للناسب لا تفرق بين السلع المختلفة بل تعاملها جميعاً نفس المعاملة وهذا عيب الأرقام القياسية التي تكون تبعاً لهذه الطريقة ، ذلك لأن السلع تختلف من حيث أهميتها وتأثير ذلك على المتوسط العام للأسعار . لذلك يحسن تكوين الأرقام القياسية بترجيح مناسيب أسعار السلع تبعاً لأهميتها . وللترجيح يمكن استخدام قيم السلع كأوزان . هذه القيم يمكن أن تحسب تبعاً للتوافق الآتية :

١ - سعر سنة الأساس × كمية سنة الأساس = ع . ك . م . مثلاً

٢ - سعر سنة المقارنة × كمية سنة المقارنة = ع . ك . م . مثلاً

٣ - سعر سنة الأساس × كمية سنة المقارنة = ع . ك . م . مثلاً

٤ - سعر سنة المقارنة × كمية سنة الأساس = ع . ك . م . مثلاً

وبذلك يمكن استخدام إحدى الصيغ الآتية لحساب الرقم القياسي تبعاً للوسط الحسابي المرجح :

$$\frac{\sum م . س}{\sum م} \quad (١)$$

$$\frac{\text{كـ س } ١٢}{١٢ \text{ كـ}} - ٢$$

$$\frac{\text{كـ س } ٢٢}{٢٢ \text{ كـ}} - ٣$$

$$\frac{\text{كـ س } ٣٢}{٣٢ \text{ كـ}} - ٤$$

وبالمثل نحصل على الوسط التوافقي المرجح بهذه الاوزان الاربعة كالآتي :

$$\frac{\text{كـ } ٠.٢}{\left(\frac{٠.٢}{\text{س}}\right) \text{ كـ}} - ١$$

$$\frac{\text{كـ } ١٢}{\left(\frac{١٢}{\text{س}}\right) \text{ كـ}} - ٢$$

$$\frac{\text{كـ } ٢٢}{\left(\frac{٢٢}{\text{س}}\right) \text{ كـ}} - ٣$$

$$\frac{\text{كـ } ٣٢}{\left(\frac{٣٢}{\text{س}}\right) \text{ كـ}} - ٤$$

كذلك يمكن أن نحصل على الوسط الهندسي المرجح بهذه الاوزان الاربعة كالآتي :

وإذا أردنا مقارنة الاسعار في فترة معينة تبعد عنها بضع فترات وكان لدينا الارقام القياسية المحسوبة على الاساس المتحرك فما علينا إلا ضرب الارقام المتتالية حتى نصل إلى الفترة المطلوبة ثم قسمة حاصل الضرب على ١٠٠. ١-١-١ حيث ن عدد الارقام المضروبة في بعضها البعض .

مثال ٦٨ :

إذا رمزنا للرقم القياسي لسنة ١٩٦٤ بالنسبة لسنة ١٩٦٣ بالرمز r_1
 r_1 » ١٩٦٤ » » ١٩٦٥ » » »
 r_2 » ١٩٦٥ » » ١٩٦٦ » » »
ثم أردنا إيجاد الرقم القياسي في سنة ١٩٦٦ بالنسبة لسنة ١٩٦٣ يكون كالآتي :

$$\frac{r_1 \times r_2 \times r_3}{100 \times 100}$$

ونلاحظ ان النتائج التي نحصل عليها للرقم القياسي إذا استخدمنا أساساً ثابتاً أو أساساً متحركاً تختلف عن بعضها ، كذلك إذا أرجعنا رقماً قياسياً على أساس متحرك إلى أساس ثابت فإن النتيجة تختلف عما لو حسب بالرقم القياسي على الأساس الثابت .

المشكلة الخامسة :

يمكن أن يكون الترجيح مباشراً أو غير مباشر . ففي حالة الترجيح المباشر تعطى كل سلعة داخلية في تكوين الرقم القياسي وزن معين يدل على أهميتها النسبية أي على أهميتها بالنسبة للسلع الأخرى التي تدخل في تكوين الرقم . ويتوقف اختيار الأوزان على البيانات الموجودة لدينا ، فإذا كان لدينا بيانات عن كميات الاستهلاك أو الانتاج من السلع المختلفة في سنة الأساس فقط

اضطربنا إلى استخدام هذه الكميات بالرغم من أن سنة الأساس قد تكون بعيدة نوعاً ما ، الأمر الذي يجعل الأهمية النسبية للسلع المختلفة قد تغيرت في الفترة الطويلة بين سنة الأساس وسنة المقارنة . وإذا وجد لدينا بيانات عن هذه الكميات في سنة المقارنة فيكون من الأفضل استخدامها كأوزان حيث أنها تعبر تعبيراً دقيقاً عن الأهمية الحاضرة للسلع المختلفة ، وإن كنا نشك كثيراً في إمكان وجود بيانات عن هذه الكميات في الوقت الذي نرصد فيه تكوين الرقم القياسي حيث أن الغالب أن هذا النوع من الكميات لا ينشر بيان عنها إلا بعد الفترة التي تتعلق بها .

وبالنسبة للرقم القياسي لنفقة المعيشة فإن الأوزان تحسب على أساس القيام ببحث استقصائي عن ميزانية الأسرة نتوصل منه إلى معرفة متوسط نسبة ما ينفق من ميزانية الأسرة على السلع المختلفة التي تدخل في معيشتها .

أما الترجيح غير المباشر يكون يحمل عدد السلع والأنواع التي تدخل في كل مجموعة من المجموعات التي يتكون منها الرقم القياسي يتناسب مع أهمية المجموعة ، ذلك لأن كثرة عدد السلع المأخوذة من أي مجموعة تعمل على تعزيز هذه المجموعة وترجيح التغيرات التي تحصل في أسعار سلعها وبالعكس فإن قلة عدد السلع المأخوذة في أي مجموعة ينقص من أهمية التغيرات التي تحصل في أسعار سلعها . كذلك يمكن أن يكون الترجيح الغير مباشر على أساس السلع وليس على أساس المجموعات وذلك بأن نأخذ عدة أنواع لنفس السلعة وندخل أسعار هذه الأنواع في تكوين الرقم القياسي وكأننا بذلك رجحنا هذه السلعة ترجيحاً غير مباشر وعززنا أهميتها بالنسبة إلى غيرها .

وكما أشرنا سابقاً لا نكون في حاجة إلى الدقة المتناهية في حساب الأوزان ، ذلك لأنها تضرب في البسيط وفي المقام وبذلك لا يؤثر الخطأ البسيط فيها على النتيجة تأثيراً كبيراً . ولهذا يحسن دائماً أن نأخذ أعداد صحيحة ونتغاضى عن الكسور حتى لا نعقد كثيراً من العمليات الحسابية .

المشكلة السادسة :

عند وضع خطة العمل لتركيب رقم قياسي يجب أن نلاحظ اننا نحتاج الى نشر النتيجة في الوقت المناسب كي يمكن الاستفادة منها فائدة عملية فلا تقتصر فائدتها للذكرى والتاريخ فقط ، وإذا تذكرنا ان بعض الارقام القياسية تنشر كل شهر أو ثلاثة أشهر أدركنا أهمية وضع الخطة المناسبة لتحقيق هذه الغاية ، ولهذا يحسن أن تكون هناك جداول معينة مدوّنة بها كل البيانات المطلوبة عدا البيانات الخاصة بالأسعار في فترة المقارنة التي يجب أن يוכל جمعها الى موظفين مدربين بحيث يدركون أهمية جمع هذه البيانات في نفس الوقت وبدقة تامة وعن نفس السلع المحددة ، وأحب أن أنبه الى ان الاعتماد على النشرات الرسمية للأسعار لا يعطينا فكرة دقيقة إلا إذا تأكدنا ان التسعيرة الرسمية يتبعها التجار بدقة تامة ، ولكن الغالب ان تختلف الاسعار الواقعية اختلافاً كبيراً من الاسعار المحددة من قبل الحكومة ، ولهذا يحسن عدم الاعتماد عليها وأن تقوم الادارة الاحصائية المختصة بجمع الاسعار التي يجري التعامل على أساسها في الاسواق المختلفة خاصة اذا تذكرنا ان الاسعار الرسمية تكون غالباً محددة لجميع أنواع الأسواق على السواء ، ولكن الواقع ان الاسعار تختلف تبعاً لما إذا كان البائع في سوق القرية أو في سوق المدينة وتبعاً لما إذا كان البائع عادياً أو مؤسسة تعاونية وكذلك تبعاً للحي الذي يوجد به البائع ، الوقت الذي تجمع فيه الاسعار .

ويلاحظ البعض ان قيام موظفو الادارة الاحصائية بجمع الاسعار قد يؤدي إلى عدم دقة البيانات الخاصة بها ، خاصة إذا كان هناك نظام للتسعيرة الجبرية تقوم الحكومة بفرضه على الاسواق . ولهذا يرون أن يוכל جمع الاسعار الى لجنة من ربات المنازل غير معروفات لدى التجار حتى تستطيع الادارة الاحصائية أن تطمئن إلى البيانات المجموعة .

وهناك اعتبار آخر يجب ملاحظته ، ذلك اننا لا يجب أن نتمسك بصيغة

معينة للرقم القياسي لأننا نعتقد انها أفضل من ناحية الدقة ، إذ يجب أن نأخذ في اعتبارنا امكانية الحصول على البيانات اللازمة لحساب الرقم القياسي تبعاً لهذه الصيغة ، وهذا غالباً يتوقف على تقدم الوعي الاحصائي في الدولة عامة ولدى التجار على الوجه الاخص . ففي مصر مثلاً من السهولة أن تجمع بيانات عن الاسعار ولكن من الصعب معرفة بيانات عن الكميات المباعة أو المنتجة ، ذلك لأن التاجر نفسه لا يعرف ذلك إلا بعد الرجوع الى دفاتره إذا كانت لديه دفاتر منتظمة . هذا فضلاً عن ان هذا البيان بالذات يعتبره التاجر سرّاً من أسرارهِ وليس عنده الاستعداد للدلاء به للغير . في مثل هذه الاحوال يجب استبعاد جميع الصيغ التي تدخل فيها ك لأن تمسكنا بواحدة من هذه الصيغ سيضيع وقتاً ، فضلاً عن اننا نعرف تماماً ان البيانات التي نجمعها ليست دقيقة .

اختبار الارقام القياسية :

رأينا كيف ان هناك صيغ مختلفة لتركيب الرقم القياسي والمفاضلة بينها لا بد ان تتناول الناحيتين العملية والنظرية ، وقد أشرنا من قبل الى المفاضلة على أساس الناحية العملية حيث رأينا كيف ان توفر البيانات والحاجة الى نشر النتائج في الوقت المناسب تحدد لنا الى حد بعيد الصيغة التي يحسن بنا استخدامها . أما المفاضلة على أساس الناحية النظرية فقد بحثها الاستاذ ارفنج فيشر حيث يرجع لديه الفضل الأكبر في هذا الموضوع .

اقترح فيشر عدة اختبارات بحيث إذا اجتازت احدى الصيغ هذه الاختبارات جميعها أمكن القول انها أفضل الصيغ . وقد وجد فيشر أن الصيغة التي تقوم على اساس الوسط الهندسي للرقم القياسي المرجح بكميات سنة الأساس والرقم المرجح بكميات سنة المقارنة هي أفضل الصيغ جميعاً حيث اجتازت الاختبارات التي يقترحها ولذا أطلق عليه الرقم القياسي الامثل .

اختبار ترتيب وضع السلع :

من الواضح ان النتيجة التي نحصل عليها من استخدام أي صيغة للرقم القياسي لا يجب أن تتغير إذا غيرنا من ترتيب السلع في المعادلة الجبرية سواء في البسيط أو في المقام . وفي الواقع نجد أن كل الصيغ تستوفي هذا الشرط .

اختبار الانعكاس في الزمن :

يقصد بذلك إذا كانت الصيغة المستخدمة في حساب الرقم القياسي يمكن أن تعمل في الاتجاهين تبعاً للزمن ؛ بمعنى انه إذا ارتفع سعر السكر من ليرة في سنة ١٩٦٠ وأصبح ٢ ليرة في سنة ١٩٦٥ - ارتفاع بنسبة ٢٠٠٪ - فان سعر السكر في سنة ١٩٦٠ إلى سنة ١٩٦٥ هو ٥٠٪ ، فكل من الرقمين هو مقلوب الآخر وبذلك يكون حاصل ضربها يساوي واحد صحيح . وبالمثل إذا كانت نتيجة الرقم القياسي على أساس صيغة معينة = ٢٠٠ على أساس سنة ما ، فان هذه الصيغة لا بد أن تعطينا مقلوب هذا الرقم إذا عكسناها بمعنى ان الرقم القياسي في سنة الأساس منسوباً إلى سنة المقارنة لا بد أن يكون ٥٠ . وبذلك لكي تكون الصيغة صحيحة يجب إذا ضربت في مقلوبها تعطينا ناتجاً = ١ صحيح ، أما اذا لم نحصل على هذه النتيجة فلا بد أن يكون هناك تحيز فيها والبديل الزمني لأية صيغة من صيغ الرقم القياسي هو نفس الرقم محسوباً بطريقة عكسية بحيث تعتبر سنة الأساس سنة مقارنة ، وسنة المقارنة سنة أساس اي اننا نضع ع. بدل ع. وكذلك ك. بدل ك. اذا وجدت .

$$\text{فالرقم : } \frac{1\text{ع.}}{\text{ع.}} \text{ بديلة الزمن } \frac{\text{ع.}}{\text{ع.}}$$

ونلاحظ ان ثلاث صيغ فقط تجتاز هذا الاختبار هي :

١ - الرقم التجميعي البسيط .

٢ - الرقم التجميعي المرجح بالوسط الحسابي لكميات سنّي الاساس والمقارنة

٣ - الرقم الأمثل .

(يمكن للطالب أَل يطبق الانعكاس في الزمن كي يتأكد من ذلك)

اختبار الانعكاس في المعامل :

قيمة اية سلعة في فترة ما تساوي حاصل ضرب الكمية المنتجة من هذه السلعة في سعرها وهي تساوي جبرياً ع ك . وبذلك فالسعر والكمية هما معاملا القيمة ، والرقم القياسي للقيمة في سنة ما بالنسبة الى سنة اخرى يساوي جبرياً
$$\frac{ع.ك}{ع.ك}$$

على هذا الاساس إذا كوّننا رقماً قياسياً لسعر سلعة ما (منسوب السعر) وكوّننا كذلك رقماً قياسياً (منسوب) لكمية هذه السلعة في نفس المدة ، فمن المنطق أن يكون حاصل ضرب هذين الرقمين يساوي النسبة بين قيمتي هذه السلعة وإلا كان الرقم القياسي المستعمل بالنسبة للسعر والكمية خاطئاً في تصويره للتغير الذي يحدث في هاتين الظاهرتين .

وقياساً على ذلك إذا استعملنا أي صيغة من الصيغ التي نعرفها للأرقام القياسية في حساب الرقم القياسي للأسعار والرقم القياسي للكميات يجب أن يكون حاصل ضرب هذين الرقمين مساوياً للرقم القياسي لقيم هذه السلع في نفس المدة وإلا كانت الصيغة المستعملة خاطئة في تصويرها للتغير الذي يحدث في هاتين الظاهرتين .

هذا الاختبار يسمى الانعكاس في المعامل أو الانعكاس المعاملي ويتلخص

في أن أي صيغة للرقم القياسي إذا ضربت في بديلها المعاملي كانت النتيجة رقماً قياسياً للقيمة ، والبديل المعاملي لأي صيغة هو نفس الصيغة موضوعاً فيها ع بدل ك ، و ك بدل ع أي الكمية بدل السعر والسعر بدل الكمية . بمعنى أنه إذا كان لدينا رقماً قياسياً للأسعار مرجحاً بالكميات فإن البديل المعاملي يكون رقماً قياسياً للكميات مرجحاً بالأسعار . ويمكن وضع هذا الشرط في الصيغة الجبرية الآتية :

$$\text{ان الرقم القياسي} \times \text{البديل المعاملي له} = \frac{\text{م.ع.ك.}}{\text{م.ع.ك.}}$$

وبتطبيق هذا الاختبار يتضح لنا أن الرقم الأمثل هو وحده الذي يحتاج وهو لهذا السبب أحسن الأرقام القياسية وأصحها من الناحية النظرية ومن ثم كانت تسميته بالرقم الأمثل .

ويمكن جبرياً إثبات انطباق هذا الشرط على الرقم الأمثل كالآتي :

$$\sqrt{\frac{\text{م.ع.ك.}}{\text{م.ع.ك.}} \times \frac{\text{م.ع.ك.}}{\text{م.ع.ك.}}} = \text{الرقم الأمثل}$$

$$\sqrt{\frac{\text{م.ك.ع.}}{\text{م.ك.ع.}} \times \frac{\text{م.ك.ع.}}{\text{م.ك.ع.}}} = \text{البديل المعاملي}$$

بضرب الرقم الأمثل في البديل المعاملي له نحصل على :

$$\frac{\cancel{\text{م.ع.ك.}} \times \cancel{\text{م.ع.ك.}} \times \cancel{\text{م.ع.ك.}} \times \cancel{\text{م.ع.ك.}}}{\cancel{\text{م.ع.ك.}} \times \cancel{\text{م.ع.ك.}} \times \cancel{\text{م.ع.ك.}} \times \cancel{\text{م.ع.ك.}}} \sqrt{\quad}$$

$$\frac{(\text{م.ع.ك.})^2}{(\text{م.ع.ك.})^2} \sqrt{\quad}$$

$$= \frac{\text{م.ع.ك.}}{\text{م.ع.ك.}} \text{ وهذا هو الرقم القياسي للقيمة.}$$

تعديل الأرقام القياسية :

اتضح لنا أن هذه الاختبارات كانت قاسية بحيث لم تنطبق كل الشروط التي اقترحها فيشر إلا على الرقم الأمثل فقط . ولما كان من الصعب أحياناً استخدام الرقم الأمثل نظراً لاحتمال عدم توفر البيانات اللازمة لتركيبه ، لهذا نحاول تعديل الأرقام القياسية التي لا تنعكس في الزمن أو في المعامل بحيث تصبح قابلة للانعكاس ، ولكي نتوصل إلى ذلك نبحث عن معكوس الرقم .

والمعكوس الزمني للرقم هو خارج قسمة الواحد الصحيح على بديله الزمني فمثلاً الرقم التجميعي المرجح بكميات الأساس لا ينعكس في الزمن ومعادلته هي :

$$\frac{\text{م.ع.ك.}}{\text{م.ع.ك.}} \quad \text{وبديله الزمني} \quad \frac{\text{م.ع.ك.}}{\text{م.ع.ك.}}$$

$$\text{وبذلك يكون المقلوب الزمني} \quad \frac{\text{م.ع.ك.}}{\text{م.ع.ك.}}$$

وبذلك يكون الوسط الهندسي للرقم الأصلي ومقلوبه الزمني يقبل الانعكاس في الزمن . هذا الوسط الهندسي هو الرقم القياسي المعدل . وهكذا يمكننا تعديل أي رقم قياسي لكي ينعكس في الزمن بأن نوجد الوسط الهندسي بينه وبين مقلوبه الزمني .

كذلك يمكن تعديل الرقم القياسي حتى ينعكس في المعامل وذلك بإيجاد الوسط الهندسي بين الرقم ومقلوبه المعاملي . والمقلوب المعاملي لأي رقم هو خارج قسمة منسوب القيم $\frac{\text{م.ع. ك.}}{\text{م.ع. ك.}}$ على البديل المعاملي للرقم نفسه .

فمثلاً الرقم المرجح بكميات سنة الأساس $\frac{\text{م.ع. ك.}}{\text{م.ع. ك.}}$

بديله المعاملي $\frac{\text{م.ع. ك.}}{\text{م.ع. ك.}}$

المقلوب المعاملي = $\frac{\text{م.ع. ك.}}{\text{م.ع. ك.}} \div \frac{\text{م.ع. ك.}}{\text{م.ع. ك.}} = \frac{\text{م.ع. ك.}}{\text{م.ع. ك.}}$

ويكون الوسط الهندسي للرقم الأصلي ومقلوبه المعاملي $\sqrt{\frac{\text{م.ع. ك.}}{\text{م.ع. ك.}} \times \frac{\text{م.ع. ك.}}{\text{م.ع. ك.}}}$

وهذا هو الرقم الأمثل الذي نعرف انه ينعكس في المعامل .

وإذا كان الرقم لا ينعكس في الزمن أو في المعامل فيجب معالجته مرتين الأولى بالنسبة إلى الزمن والثانية بالنسبة إلى المعامل وسواء بدأنا بهذا أو بذلك تكون النتيجة رقماً يقبل الانعكاس في الزمن وفي المعامل .

ولكن هذا التعديل - مع الأسف - يعطينا في جميع الأحوال شيئاً

أكثر تعقيداً في التركيب وأشد ارهاقاً في الحساب من الصيغ الأصلية المقصود تعديلها وهذا التعقيد يقلل من فائدة هذا التعديل وأهميته من الناحية العملية .

الرقم القياسي للكميات :

يمكن حساب رقماً قياسياً للكميات بنفس الطرق السابقة مع ملاحظة أن تؤخذ الأسعار أو القيم كأوزان :

$$\text{مثلاً} \quad \frac{\text{م. ك. ١ ع.}}{\text{م. ك. ع.}} \times 100 \text{ رقم قياسي للكمية مرجحاً بأسعار الأساس .}$$

$$\text{أو} \quad \frac{\text{م. ك. ١ ع.}}{\text{م. ك. ع.}} \times 100 \text{ رقم قياسي للكمية مرجحاً بأسعار المقارنة .}$$

وهكذا بالنسبة لباقي الصيغ التي أشرنا إليها بالنسبة للأرقام القياسية للأسعار .

كذلك يمكن حساب رقم قياس للكميات بطريقة المناسبة حيث يحسب

منسوب الكمية لكل متغير على حده ($\frac{1}{100} \times 100$) ثم بإيجاد الوسط الحسابي أو الوسط الهندسي أو الوسط التوافقي (بسيطاً أو مرجحاً) ،
للمناسيب المتغيرات التي نقرر ادخالها في الرقم القياسي نحصل على الرقم الذي نريده وهو مثل الرقم القياسي للأسعار يقيس التغير النسبي في كميات المتغيرات بالنسبة لفترة الأساس ويفضل الوسط الهندسي للمناسيب خاصة في الرقم القياسي لأسعار الجملة حيث انه يشمل عدداً كبيراً من السلع التي تختلف

اسعارها كثيراً ، وميزة الوسط الهندسي انه يقلل من تأثير القيم الكبيرة جداً والقيم الصغيرة جداً ، وذلك بعكس الوسط الحسابي الذي يتأثر بها كثيراً ، وفيما عدا الرقم القياسي لأسعار الجملة لا يستخدم الوسط الهندسي نظراً للعمليات الحسابية المعقدة التي يحتاجها خاصة إذا كان الرقم الذي نحسبه مرجحاً . وبذلك يكون الوسط الحسابي للمناسيب هو الوسط الشائع استعماله في حساب الارقام المختلفة .

استبعاد اثر تغيرات الاسعار :

بالاضافة إلى فائدة الرقم القياسي للأسعار في دراسة التغيرات النسبية في الاسعار على مر الزمن ، فإنه يمكن بواسطة هذه الارقام معرفة التغيرات الحقيقية التي تحدث في الظواهر التي تتعرض قيمها لتضخم شديد بسبب ارتفاع الاسعار مثل الدخل القومي والاجور وقيم الصادرات والواردات ... الخ . ان التضخم الشديد في قيم هذه الظواهر يخفى تغيراتها الحقيقية التي قد نكون مهتمين بالبحث عنها . وبواسطة الارقام القياسية للأسعار نستطيع استبعاد اثر تغيرات الاسعار وبذلك لا يبقى في الظاهرة من تغير الا التغيرات الحقيقية أي تغيرات الكميات الخاصة بمفردات الظاهرة والتي من الواضح لا يمكن جمعها سوا بسبب اختلاف وحداتها ويكون هذا الاستبعاد تطبيقاً للقاعدة ، الكمية \times السعر = القيمة ، وبذلك فان قسمة القيمة على الرقم القياسي للأسعار يعطينا رقماً يمثل التغير الحقيقي ، أي التغير في كمية الظاهرة أو الظواهر موضوع البحث . فاذا قسمنا قيمة الدخل القومي على الرقم القياسي للأسعار نحصل على رقم معين يبين قيمة الدخل القومي مقاسة بأسعار سنة الأساس . كذلك اذا قسمنا الرقم القياسي لكسب العمل على الرقم القياسي لنفقة المعيشة نحصل على رقم يمثل التغير النسبي في مستوى المعيشة (من الناحية المادية فقط) والذي نطلق عليه أحياناً الرقم القياسي للاجر الحقيقي .

الارقام القياسية الشائعة

الرقم القياسي لنفقة المعيشة :

يعتقد الاحصائيون في الوقت الحاضر ان تسمية هذا الرقم هكذا تسمية مضللة . ويفضلون تسميته بالرقم القياسي لاسعار التجزئة . ويقيس هذا الرقم تكلفة شراء مجموعة معينة من السلع التي يُعتقد انها تمثل استهلاك الطبقة التي من أجلها حسب هذا الرقم . ويحسب هذا الرقم بصيغة مرجحة بأوزان سنة الاساس . وتعتمد الأوزان التي تستخدم في الترجيح على نتائج بحث ميزانية الاسرة . والطريقة الشائعة في حساب هذا الرقم هي على اساس الوسط الحسابي المرجح لمنايب الاسعار .

وفي الغالب يحسب هذا الرقم شهرياً ، حيث تجمع الاسعار في يوم معين في منتصف الشهر . وفي بعض الدول تجمع الاسعار اسبوعياً ثم يحسب المتوسط لها ، وتكون نتيجة الرقم بذلك اكثر دقة ، إلا ان ذلك يتطلب جهداً أكبر ، خاصة اذا كان الرقم يتضمن عدداً كبيراً من السلع .

وتقوم مصلحة الاحصاء في أغلب الدول بجمع الاسعار اما بارسال موظفيها إلى محلات معينة او بمطالبة اصحاب هذه المحلات بالتبليغ عن اسعار السلع المحددة مرة كل شهر . والمشكلة الاساسية في هذه العملية هي ضرورة التأكد من ان الاسعار التي تجمعها المصلحة تمثل اسعار نفس السلع كل شهر .

الرقم القياسي لاسعار الجملة :

يقيس هذا الرقم التغيرات النسبية في اسعار السلع التي يتم التبادل فيها في اسواق الجملة . وكان الاحصائيون يأملون في بادئ الأمر الحصول على مقياس يمكن ان يقيس التغيرات العامة في الاسعار ، إلا انه في الوقت الحاضر أصبح التركيز على مقياس لتغيرات اسعار مجموعات معينة من السلع .

ويحسب هذا الرقم بصيغة مرجحة ، حيث تستخدم قيم السلع التي جرى تبادلها في فترة الأساس . إلا انه في بعض الدول يجري ترجيح أسعار السلع بأخذ أسعار عدد معين من أصناف كل سلعة يتناسب مع أهميتها ، بمعنى انه كلما أردنا إعطاء السلع وزناً أكبر ندخل في الرقم عدداً أكبر من أصنافها .

وفي الغالب يحسب الرقم القياسي لأسعار الجملة بطريقة الوسط الهندسي المرجح للناسيب خاصة إذا كنا نتبع نظاماً بسيطاً للترجيح (مثلاً أخذ أسعار عدد معين من أصناف كل سلعة) .

الرقم القياسي للانتاج الصناعي :

هذا رقم قياسي الكميات هدف إلى قياس التغيرات في انتاج مجموعة معينة من الصناعات . وفترة الأساس لهذا الرقم هي غالباً سنة ، والأوزان هي القيمة الصافية لانتاج كل سلعة في فترة الأساس . ويحسب الرقم بطريقة الوسط الحسابي المرجح للناسيب الكميات .

وتجمع المعلومات الخاصة بهذا الرقم بمطالبة عينة من المؤسسات بالتبليغ شهرياً عن إنتاجها . وفي بعض الحالات التي لا يمكن فيها جمع بيانات عن الإنتاج تؤخذ معلومات أخرى تكون مؤشر على الانتاج وذلك مثل البيانات الخاصة بالتوظيف ، على ان ذلك لا يجب أن يتبع الا اذا يثسنا من الحصول على بيانات خاصة بالإنتاج نفسه حيث ان بيانات التوظيف تجعل الرقم لا يصلح كمقياس الإنتاجية .

الارقام القياسية لحجم الصادرات والواردات .

الهدف من هذه الارقام هو قياس التغيرات في قيم الواردات والصادرات بأسعار ثابتة . ان التغير في قيم الواردات او الصادرات يمكن أن يكون نتيجة تغير كميات السلع المستوردة أو المصدرة أو نتيجة تغيرات اسعارها أو

نتيجتها سوية . والهدف من هذه الارقام القياسية هو استبعاد اثر تغيرات الاسعار حتى يمكن ملاحظة التغيرات الحقيقية في حجم التجارة .

ويحسب هذا الرقم بطريقة الوسط الحسابي المرجح لمناسيب الكميات ، حيث تستخدم قيم السلع في فترة الاساس كأوزان ، وفترة الاساس هي غالباً سنة . وتحسب المناسيب على أساس المعلومات التي تظهرها النشرات الاحصائية للتجارة الخارجية حيث نجد فيها بيانات عن قيم وكميات السلع المستوردة او المصدرة كل شهر . وفي بعض الحالات عندما لا تتوفر بيانات دقيقة ومناسبة عن الكميات يمكن التوصل الى رقم قياسي للكميات بقسمة رقم القيم على رقم الاسعار الخاصة بالبضائع التي جرى التبادل فيها والذي يحسب خصيصاً لهذا الغرض . وفي هذه الحالة يكون رقم الكميات الناتج رقماً مرجحاً بقيم فترة المقارنة وليس قيم فترة الاساس . ومن الواضح أنه في هذه الحالة يحسن استخدام صيغة فيشر لحساب الرقم القياسي للأسعار .

وإذا حسبنا الرقم القياسي لأسعار الصادرات أو الواردات فإنه يظهر نسب التغيرات في قيم التجارة التي ترتبت على تغيرات الأسعار ، فإذا أظهر الرقم القياسي لأسعار الواردات ارتفاعاً بنسبة ٢٠ ٪ مثلاً فإن ٢٠ ٪ من الزيادة في قيم الواردات تكون راجعة إلى هذا التغير في أسعارها . وإذا توفر لدينا رقماً قياسياً لقيم الواردات أو الصادرات ورقماً قياسياً لكمياتها بقسمة الأول على الثاني نحصل على رقم قياسي لأسعارها .

ومن الواضح أن مقارنة الرقم القياسي لأسعار الصادرات بالرقم القياسي لأسعار الواردات تدلنا على معدل التبادل في التجارة الخارجية وهو مقياس ذو أهمية كبرى في تحديد الكسب الذي يعود على الدولة من تجارتها الخارجية .

تمارين

(١) الآتي بيانات عن انتاج احدى الدول من القمح بالطن :

السنة	١٩٥٠	١٩٥١	١٩٥٢	١٩٥٣	١٩٥٤	١٩٥٥	١٩٥٦
الانتاج	١٠١٩	٩٨٨	١٣٠٦	١١٧٣	٩٩٤	٩٣٥	١٠٠٤

احسب مناسيب الانتاج على أساس عام ١٩٥٥
 » » » » متوسط الأعوام من ١٩٥٠ - ١٩٥٣ .

(٢) إذا علمت أن منسوب الكمية لعام ١٩٥٨ على أساس عام ١٩٤٩ = ١٠٥ وان منسوب الكمية لعام ١٩٥٨ على أساس ١٩٥٣ = ١٤٠ - احسب منسوب الكمية عام ١٩٥٣ على أساس عام ١٩٤٩ .

(٣) تتوقع احدى المؤسسات زيادة مبيعاتها من سلعة ما بنسبة ٥٠٪ في العام القادم . بأي نسبة يجب أن ترفع سعر هذه السلعة حتى يمكن أن تزيد إيراداتها إلى الضعف .

(٤) إذا كانت مناسيب الأسعار في الفترة من عام ١٩٥٦ إلى عام ١٩٦٠ بطريقة السلسلة هي على التوالي ١٢٥ ، ١٢٠ ، ١٣٥ ، ١٥٠ ، ١٧٥ . حول هذه المناسيب إلى مناسيب على أساس عام ١٩٥٦ .

(٥) من البيانات الآتية احسب الرقم القياسي للأسعار بصيغة فيشر وكذلك الرقم القياسي للكميات (١٩٥٠ على أساس ١٩٤٩) .

السلع	اسعار ١٩٤٩	اسعار ١٩٥٠	كميات انتاج ١٩٤٩	كميات انتاج ١٩٥٠
ا	٣٩٥	٣٨٩	٩٦٧٥	٩٧١٧
ب	٦١٥	٦٢٢	١١٧٧	١١٥٥
ج	٣٤٥	٣٥٤	٧٧٩٣	٧٥٣٩

(٦) إذا علمت أن متوسط اجر العامل في السنوات من عام ١٩٤٧ إلى عام ١٩٥١ هو ١٠١٩ ، ١٠٣٢ ، ١٠٤٤ ، ١٠٥٧ دينار . وان الأرقام القياسية للأسعار هي على التوالي ١٠٠ ، ١٠٧٦ ، ١٠٦٦ ، ١٠٧٦ . احسب متوسط اجر العامل بأسعار عام ١٩٤٧ .

الفصل الرابع عشر

السلاسل الزمنية

مقدمة :

يتم الاقتصاديون بنوعين من الدراسات الاقتصادية ، النوع الأول هو الذي يتجه إلى تحليل حالة النشاط الاقتصادي التي يمكن ان توجد إذا كانت القوى الاقتصادية في توازن تام بحيث لا يصيب النشاط الاقتصادي أي اختلال ، هذا النوع من التحليل يطلق عليه الاقتصاديون التحليل الاستاتيكي . والنوع الثاني وهو الذي يطلق عليه التحليل الديناميكي يتجه إلى دراسة ما يمكن أن يصيب النشاط الاقتصادي من اختلال نتيجة تغير القوى الاقتصادية المختلفة فبينما نهتم في النوع الأول بدراسة الحالة التي يصل إليها النشاط الاقتصادي بعد الوصول إلى التوازن نهتم في النوع الثاني بتحليل ما يمكن أن يحدث من تغيرات قبل الوصول إلى التوازن .

والواقع أن الاقتصاديين في العصر الحديث أصبحوا يركزون اهتمامهم على النوع الثاني من التحليل الاقتصادي حيث انهم تنبهوا إلى أن حالة التوازن ليست إلا حالة نظرية لا يمكن أن يقف عندها النشاط الاقتصادي الذي يبدو أنه دائم التغير والحركة . هذا النوع من الدراسة يقوم إلى حد بعيد على الدراسة الاحصائية للسلاسل الزمنية ، تلك الدراسة التي كان للاقتصاديين الفضل الكبير في تقدمها تقدماً كان من نتيجته ابتداء الأساليب الاحصائية

التي يمكن استخدامها في تحليل ما يصيب النشاط الاقتصادي من تغيير مع الزمن ، وهي أساليب تختلف تماماً عن تلك التي تستخدم في تحليل التوزيعات التكرارية التي تدرس الظواهر في حالة السكون . ولو ان الاقتصاديين كان لهم الفضل الأكبر في الدراسة الاحصائية للسلاسل الزمنية الا ان فائدتها ليست قاصرة عليهم فهي تهم كذلك رجال الأعمال ، والباحثين الاجتماعيين والأطباء وعلماء الحياة والعاملين في ميادين الصحة العامة .

تعريف السلسلة الزمنية :

يمكننا أن نعرف السلسلة الزمنية لاية ظاهرة بانها التطور التاريخي لهذه الظاهرة ، وبمعنى آخر هي القيم أو المقادير التي تتخذها هذه الظاهرة في فترات زمنية متتابعة قد تكون أيام أو أسابيع أو أشهر أو سنين . ولتركيب سلسلة زمنية لظاهر ما نعمل أولاً على مشاهدتها مدة من الزمن ونسجل أثناء المشاهدة القيم المختلفة التي تتخذها في فترات زمنية منتظمة فيمكننا مثلاً أن نكوّن سلسلة زمنية لكمية البنكنوت المتداول في دولة ما بسؤال البنك المركزي في هذه الدولة أو بالرجوع إلى الاحصاءات التي ينشرها عن هذا الموضوع فنحصل على قيم شهرية نضعها في جدول بين به الفترات الزمنية والقيم المقابلة للبنكنوت المتداول ، هذا الجدول يوضح السلسلة الزمنية لهذه الظاهرة خلال هذه المدة من الزمن .

والهدف من الدراسة الاحصائية للسلاسل الزمنية هو الكشف عن التغيرات التي تطرأ على الظواهر التي ندرسها أثناء مدة معينة حتى يمكن معرفة أنواع هذه التغيرات وقياس كل نوع منها . فالتحليل الاحصائي لاية سلسلة زمنية يقوم أساساً على مقارنة قيم الظاهرة في فترات متتابعة حتى يمكن الكشف عما يصيبها من نمو أو ضمور . ولذلك يجب أن نلتنبه الى ان المقارنة لكي تكون صحيحة يجب أن تقاس الظاهرة بنفس الوحدات وببنفس الطريقة في الفترات

الزمنية المتتابعة بمعنى أن لا تتغير طبيعتها أو صفاتها التي تميزها أو على الأقل لا تتغير بدرجة كبيرة .

الدراسة الاحصائية للسلسلة الزمنية :

تبدأ الدراسة الاحصائية لأي سلسلة زمنية برسم خط بياني لها يوضح تغيرها مع الزمن ، فنبين الزمن على المحور الأفقي والقيم أو المقادير على المحور الرأسي ، ثم نضع نقط أمام كل فترة زمنية والقيمة المقابلة لها ، ثم نصل هذه النقط فيتكون لدينا خط بياني يسمى أحيانا بالمنحنى التاريخي للظاهرة التي ندرسها .

بعد ذلك نتجه الدراسة الإحصائية للسلاسل الزمنية الى تحليل التغيرات التي تطرأ على الظواهر موضوع البحث إلى أنواعها المختلفة ثم قياسها وتحديد اتجاهها سواء كان نحو الزيادة أو نحو النقص والاستفادة من ذلك في فهم طبيعة هذا التغيرات فيها لكي نتوصل اليه يجب ان يكون لدينا معرفة تامة بموضوع البحث .

كما ان دراسة السلاسل الزمنية تهدف أيضاً الى الاستفادة من دراستنا لماضي الظواهر في عمل تقديرات لها في المستقبل حتى يمكن أن نستعد لمواجهة ما يطرأ عليها من تغيرات فلا نفاجأ بها . كذلك يتضمن التحليل الاحصائي للسلاسل الزمنية دراسة ما يمكن ان يوجد من ارتباط بين سلسلتين أو أكثر .

أنواع التغيرات التي يمكن أن تحتويها السلسلة الزمنية :

ليس هناك اتفاق تام بين الاقتصاديين على الأنواع المختلفة للتغيرات التي يمكن أن تظهر في السلسلة الزمنية . إلا أن هناك اربع انواع رئيسية يكاد الجميع يجمع على تمييزها ودراستها وهي التغيرات الموسمية والدورية والعرضية والاتجاه العام .

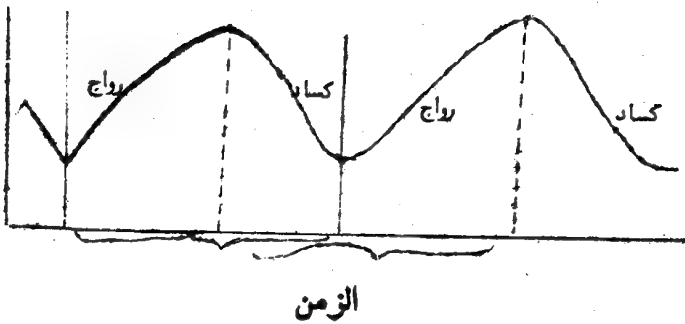
١ - التغيرات العرضية - وهي تغيرات لم يحاول الاقتصاديون وضع نظرية لتفسيرها حيث انها تنشأ عن اسباب عارضة لم تكن في الحسبان ولذا فهي شاذة وطارئة من الوجهة النظرية بمعنى انه لا يمكن التنبؤ بوقوعها أو تحديد مقدارها حيث انها لا تتبع أي قاعدة او قانون . ويترتب على ذلك اننا لا ننتظر تكرار هذا النوع من التغير وانما هو يفاجئنا . ومثل هذه التغيرات ما يصيب النشاط الاقتصادي نتيجة زلزال أو فيضان او حرب أو اضراب عام بين العمال .

على أن هذه التغيرات قد تكون من الضالة بحيث لا تستحق الانتباه اليها وقد تكون من الضخامة بحيث تثير انتباهنا ، إذ قد يترتب على هذه الضخامة نشوء نوع آخر من التغير ، فالحرب مثلاً تعتبر في بدايتها تغيراً عرضياً ولكنها بعد ذلك تصبح تغيراً دورياً له أهميته .

٢ - التغيرات الموسمية - وهي تغيرات منتظمة انتظاماً تاماً إذا قورنت بأي نوع آخر من التغيرات ، فهي تغيرات تكرر في فترات منتظمة بحيث تحدث في تواريخ معلومة في كل سنة ولا تحيد عنها بحيث يكون اتجاهها واحد لا يتغير . فكمية النقد المتداول في مصر تتجه نحو الزيادة خلال شهر سبتمبر واکتوبر ونوفمبر من كل عام وإلى النقص خلال شهر يوليو وأغسطس ، والمهم في هذه التغيرات ان أسبابها ليست كامنة في تفاعل النشاط الاقتصادي نفسه ، وانما هي أسباب مختلفة عن ذلك قد تكون التقاليد الاجتماعية مثل الأعياد وقد يكون تغير الجو مثل موسم الصيف في لبنان وقد يكون ظهور المحصول الهام في الدولة في فترة زمنية معينة مثل القطن في مصر .

وحيث أن هذه التغيرات تتصف بالتكرار فلذا يمكن دراستها احصائياً حتى يمكن قياسها ومعرفة اتجاهها والاستعداد لها في المستقبل . والدراسة الاحصائية لهذه التغيرات تقوم على أساس مشاهدة الظاهرة التي ندرسها في الفترات المتتابعة التي تتكون منها السنة مثل الأشهر او ارباع السنة .

٣ - التغيرات الدورية - وهي تغيرات منتظمة ولكن انتظامها ليس تاماً كما هو الحال في التغيرات الموسمية ، إذ أن طول الدورة الاقتصادية ليس معروفاً بدقة كما انه يتغير ولذلك لا يمكن تحديد مواعيد هذه التغيرات بدقة ، والتغيرات الدورية ليست تغيرات قصيرة الأجل مثل التغيرات الموسمية إذ لا يظهر لها أثر محسوس بين سنة وأخرى وانما تستغرق زمناً طويلاً حتى تستعيد سيرتها وقد اختلف الاقتصاديون في تحديد هذا الزمن ولكن الغالب أن الدورة تستغرق حوالي عشر سنوات ، والمقصود بطول الدورة المدة التي تمضي بين قمة الموجه والقمة التي تليها (قمة الرواج) أو بين قاع الموجه والقاع الذي يليه (قاع الكساد) كما يتضح من الرسم التالي .



وتختلف هذه التغيرات كذلك عن التغيرات الموسمية في أن أسبابها اقتصادية بحتة بمعنى أنها كامنة في تفاعل النشاط الاقتصادي نفسه وقد وضعت نظريات اقتصادية كثيرة تحاول تفسير هذه الأسباب (نظرية كينز مثلاً) .

وبالطبع يمكن دراسة هذه التغيرات دراسة احصائية لأنها كما ذكرنا تتصف بالتكرار . والدراسة الاحصائية لها تقوم على أساس مشاهدة الظاهرة التي نبحثها في مدة طويلة ، عشرين سنة مثلاً .

٤ - الاتجاه العام - أثناء حدوث الأنواع السابقة للتغيرات يطرأ على الظاهرة تغير بطيء لا يمكن ملاحظته في المدة القصيرة وإنما يكون واضحاً بعد مرور مدة طويلة من الزمن ، هذا التغير هو ما نسميه بالاتجاه العام وهو تغير طبيعي يطرأ على الظواهر المختلفة فيوضح ما يصيبها من نمو في فترة تاريخية معينة ومن ضمور في فترة تاريخية أخرى . فإذا درسنا تطور صناعة ما في دولة معينة يمكن أن نلاحظ تغيراً بطيئاً يصيبها تدريجياً بحيث تأخذ في النمو حتى تصل إلى قمة معينة ثم يطرأ عليها تغير تدريجي بطيء أيضاً في اتجاه عكسي بحيث تأخذ في الضمور بمرور الزمن . أما عن أسباب هذا النوع من التغير البطيء التدريجي بالنسبة للسائل الاقتصادي فغير واضحة تماماً وهناك آراء كثيرة في هذا الصدد .

تفسير التغيرات التي تتضمنها السلسلة الزمنية :

لا يجب أن نعتقد أن الأنواع السابقة من التغيرات تحدث منفصلة عن

بعضها بحيث يمكن القول ان الظاهرة في هذه الفترة تخضع لتغير موسمي فقط وفي تلك لتغير عرضي فقط وهكذا ، ذلك لان هذه الانواع تحدث جميعاً في وقت واحد ، فيعمل كل نوع من ناحيته على التأثير على الظاهرة بدرجة معينة وفي اتجاه معين . وبذلك يكون أي تغير يطرأ على الظاهرة هو في الواقع المحصلة لجميع القوى والمؤثرات التي تحيط بالظاهرة ، فحالة الرواج مثلاً تقوي من التغير الموسمي الذي يكون في اتجاه صعودي ، وتضعف من التغير الذي يكون في اتجاه نزولي ، والكساد على العكس يقوي من التغير الموسمي النزولي ويضعف من التغير الصعودي . والاتجاه العام نحو الصعود يضعف من الكساد ويقوي من الرواج والعكس بالنسبة للاتجاه العام نحو الضمور فانه يقوي من الكساد ويضعف من الرواج .

ولذلك تنشأ لدينا عند الدراسة الاحصائية للسلاسل الزمنية مشكلة تحليل التغير الذي يطرأ على الظاهرة إلى أنواعه الأربعة وبمعنى آخر مشكلة الفصل بينها حتى يمكن تحديد اتجاه كل منها وقياسه .

قياس التغيرات وتحديد اتجاهها :

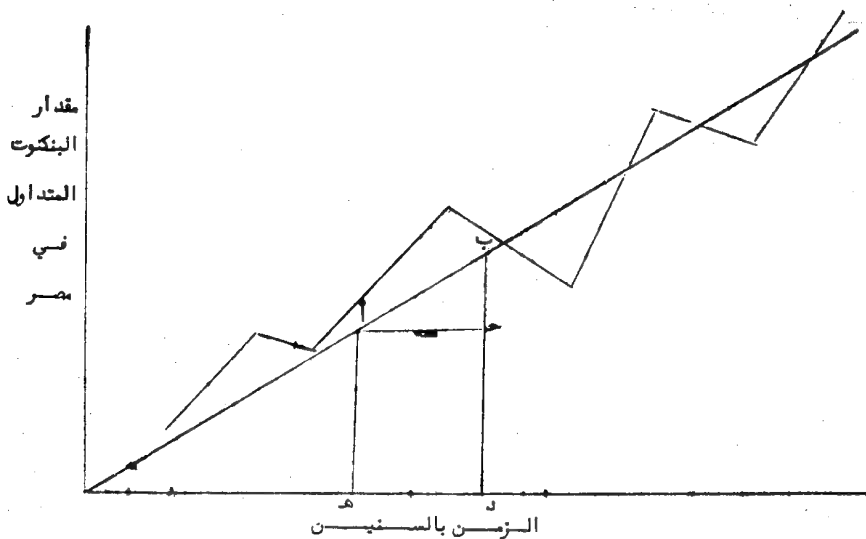
سبق ان ذكرنا أن التغيرات العرضية تحدث دون قاعدة أو قانون ولذا لا يمكن قياسها أو تقديرها حتى تحصل فعلاً . ولكن لكي نستطيع أن نقيس الانواع الأخرى من التغيرات يجب أن نعمل على استبعاد التغيرات العرضية وبتلصنا لنا ذلك بمشاهدة الظاهرة خلال مدة طويلة من الزمن حتى يمكن لهذا النوع من التغيرات أن يمحو بعضه بعضاً فتنكشف لنا الاتجاهات الحقيقية للظاهرة .

ولقياس الاتجاه العام نحسب مقدار التغير الذي يصيب الظاهرة أثناء وحدة الزمن في المتوسط أي بعد استبعاد تأثير التغيرات العرضية ، ذلك لاننا نقصد بالاتجاه العام معدل التغير في المتوسط بالنسبة لوحدة الزمن ،

ويجب ذكر عبارة « في المتوسط » حيث نقصد بها اننا استبعدنا أثر التغيرات العرضية والموسمية والدورية .

ويستبعد أثر التغيرات الموسمية بأن نأخذ القيم السنوية للظاهرة - مجموعها أو متوسطها - لأن هذه التغيرات تستعيد سيرتها كل ١٢ شهراً . أما أثر التغيرات العرضية فيستبعد بتمهيد الخط البياني الذي يمثل السلسلة الزمنية للظاهرة أو بطريقة المتوسطات المتحركة التي سيأتي الكلام عنها فيما بعد . وتمهيد الخط البياني للسلسلة يجعلنا نحصل على خط يسمى بخط الاتجاه العام وهو إما أن يكون مستقيماً أو غير مستقيم . فإذا كان مستقيماً دل ذلك على ان معدل التغير يسير بانتظام خلال الفترة التي ندرسها ، وإذا كان غير مستقيم دل ذلك على ان معدل التغير (الاتجاه العام) ليس منتظماً خلال هذه الفترة بمعنى انه قد يكون صعودياً في بادئ الامر ثم يغير من اتجاهه فيصبح نزولياً بعد مرور مدة من الزمن .

بعد تمهيد الخط البياني للسلسلة نحدد الفترة التي نريد قياس الاتجاه العام فيها ثم نحدد نقطتين على خط الاتجاه العام بحيث يقعاً خلال هذه الفترة ومنها نسقط عمودين على المحور الأفقي ومن النقطة الأولى نسقط عموداً موازياً للمحور الأفقي . ثم نحسب ظل الزاوية التي يصنعها خط الاتجاه العام مع هذا المحور فيكون هو معدل التغير الذي يميز تبعاً لتوحدات التي قيست بها الظاهرة والذي يدل على مقدار التغير في المتوسط في وحدة الزمن .



واضح من الرسم كيف مهدنا خط الاتجاه العام بحيث يكون خالياً من التعريجات غير المنتظمة أو الذبذبات قصيرة الاجل بحيث يصور الاتجاه العام للظاهرة التي ندرسها تصويراً دقيقاً على هذا الخط حددنا نقطتين : أ ، ب ومنها أسقطنا عمودين أ هـ ، ب د . ومن النقطة أ أسقطنا عموداً د على

ب د . ومعدل التغير يكون بذلك هو ظل الزاوية ب أ ح ويساوي $\frac{ب د}{أ ح}$

وحيث ان خط الاتجاه العام مستقيم فلا يتغير معدل التغير مهما غيرنا من النقط التي نحددها على هذا الخط .

وبنفس الطريقة يمكننا تمهيد خط غير مستقيم ، إلا أنه في هذه الحالة يكون هناك اتجاهان أحدهما صعودياً والآخر نزولياً فيكون ظل الزاوية في الحالة الأولى موجباً أي يكون معدل التغير بالزيادة بينما يكون ظل الزاوية في الحالة الثانية سالباً أي يكون معدل التغير بالنقص . ونلاحظ في حالة الاتجاه غير المستقيم أن الفترة الزمنية التي نأخذها لحساب معدل التغير يجب

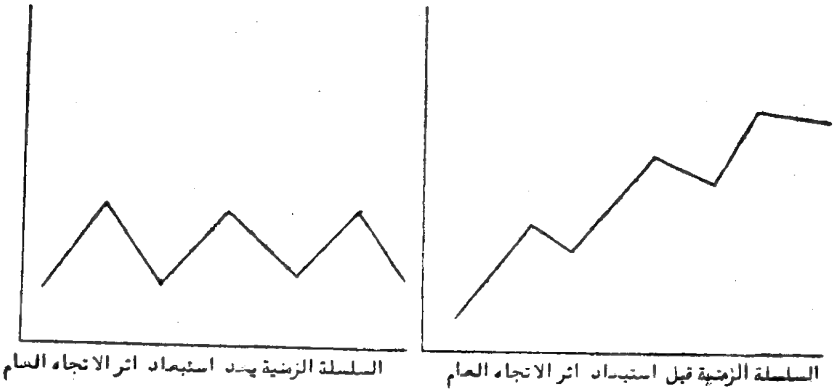
ان تكون فترة قصيرة جداً خوفاً من أن يتغير الاتجاه بمقدار كبير إذا أخذنا فترة زمنية طويلة .

على ان معدل التغير الذي نحسبه بالطريقة السابقة يكون عرضة للحكم الشخصي الذي يتوقف الى حد بعيد على رسم خط الاتجاه العام وليس من السهل أن يتفق أكثر من شخص واحد في رسم الاتجاه العام لسلسلة زمنية فيهاذبذبات كثيرة وعنيفة . ولهذا نلجأ إلى طريقة أخرى كي نضعف من التمريجات التي تتخلل المنحنى التاريخي بحيث يصبح هذا المنحنى قريباً من التمهيد فيكون من السهل بعد ذلك رسم خط الاتجاه العام . وتقوم هذه الطريقة على أساس المتوسطات المتحركة وذلك بأن نأخذ متوسطات القيم خلال عدة فترات ونعتبر هذا المتوسط كأنه قيمة الظاهرة في منتصف هذه المدة ثم نترك الفترة الأولى ونأخذ فترة أخرى من أسفل ونحسب المتوسط وهكذا ، فيصبح لدينا قيم ليس فيها أثر للتغيرات العرضية القصيرة الأجل أي خالية من الذبذبات العنيفة التي قد تكون في القيم الأصلية وبذلك يكون من السهل تحديد خط الاتجاه العام لها ، الذي نحسب على أساسه معدل التغير أي الاتجاه العام للظاهرة بنفس الطريقة السابقة .

ونلاحظ أنه كلما طالت المدة التي نحسب على أساسها المتوسطات المتحركة كلما صارت السلسلة الزمنية أكثر تمهيداً . على ان ذلك يؤدي إلى قصر السلسلة الزمنية الجديدة التي تتكون من المتوسطات المتحركة وهذا عيب إذ ان ذلك قد يؤدي الى ضياع معالم السلسلة الزمنية الأصلية حيث يصبح اتجاهها العام مختلفاً بعض الشيء عن الاتجاه العام للمتوسطات المتحركة . كذلك نلاحظ انه إذا كانت السلسلة الزمنية تتضمن تغيرات موسمية فلكي نخلص السلسلة من هذه التغيرات يجب أن تحسب المتوسطات المتحركة على أساس طول الفترة التي يحدث فيها التغير الموسمي وهو ١٢ شهراً او الاجزاء التي تتكون منها السنة والتي تكون السلسلة الزمنية الأصلية مقاسة تبعاً لها ..

وهناك طريقة ثالثة يمكننا ان نلجأ اليها لتمهيد السلسلة الزمنية وذلك بتوفيق المنحني الذي يمثل هذه السلسلة اي بتوفيق المعادلة الرياضية التي تمثل العلامة بين قيم الظاهرة والزمن . والزمن هنا يعتبر المتغير المستقل وقيم الظاهرة هي المتغير التابع . والتمهيد بهذه الطريقة ادق من التوفيق بالطريقتين السابقتين ولكن العيب الذي يؤخذ على هذه الطريقة هو صعوبتها من الناحية العملية خاصة اذا كان عدد القيم كبيراً او اذا كانت المعادلة التي نوفقها تمثل العلاقة بين الزمن والظاهرة خلال المدة المشمولة بالبيانات التي لدينا فقط بمعنى انه لا يمكن تطبيق هذه المعادلة على الماضي او المستقبل الا اذا فرضنا ان الظروف التي احاطت او ستحيط بهذه الظاهرة لا تختلف كثيراً عن الظروف التي احاطت بها اثناء المدة التي درسناها .

بعد حساب الاتجاه العام اي معدل التغير يمكن بطرحه طرحاً جبرياً اي بالإحتفاظ بإشارته الجبرية من القيم الاصلية للظاهرة ان نحصل على قيم جديدة خالية من تأثير هذا الاتجاه العام ولكنها تتضمن طبعاً الانواع الاخرى من التغيرات ويمكن تصوير هذه العملية بالرسم البياني الآتي :



وبذلك يمكننا ان نحدد الطرق الآتية لتقدير الاتجاه العام لظاهرة ما ،

اي لتقدير متوسط التغير في الظاهرة في وحدة الزمن .

١ - تحديد خط الاتجاه العام برسمه على السلسلة ثم حساب ميل الخط من الرسم البياني . ومن الواضح ان لهذه الطريقة عيب هام حيث ان النتيجة تتوقف على قدرة الانسان على تحديد الاتجاه العام للسلسلة بدقة .

٢ - تمهيد الذبذبات التي تظهر في السلسلة بحساب المتوسطات المتحركة التي يمكن بواسطتها استبعاد اثر التغيرات وبرسم السلسلة من واقع هذه المتوسطات يكون من السهل تحديد خط الاتجاه العام لها ثم حساب ميله الذي يدل على قيمة الاتجاه العام . وهذه الطريقة وان كانت اكثر دقة من الطريقة السابقة الا انها لا زالت تعتمد على حرية الانسان في رسم خط الاتجاه العام .

٣ - تمهيد السلسلة بإيجاد معادلة خط الاتجاه العام سواء كان مستقيما او منحنى وذلك باتباع طريقة المربعات الصغرى . وفي تطبيق هذه الطريقة يمكن اعتبار الفترة الاولى هي نقطة الاصل في الرسم وبذلك يعبر عنها بالرقم صفر ثم الفترة التالية بالرقم ١ وهكذا ، كما يمكن اخذ فترة متوسطة واعتبارها هي فترة الاساس اي نقطة الاصل ويعبر عنها بالرقم صفر وعن الفترات السابقة لها بالارقام ١ - ، ٢ - ، ٣ - وهكذا ابتداء من نقطة الاصل وعن التغيرات اللاحقة لها بالارقام ١ + ، ٢ + ، ٣ + وهكذا ابتداء من نقطة الاصل . واتباع الطريقة الثانية يكون افضل حيث أنه يسهل علينا العمليات الحسابية خاصة اذا كان عدد الفترات فرديا . ونلاحظ ان قيمة الاتجاه العام وهي ميل الخط الذي يمثله معادلة الاتجاه العام (القيمة م في المعادلة) ، لا تختلف سواء اتبعنا الطريقة الاولى او الثانية ، الا ان قيمة ب تتغير ولذلك بعد ايجاد المعادلة يجب ان نتبعها بملاحظة نحدد فيها السنة التي اعتبرت نقطة الاصل للمعادلة .

٤ - يمكن ان نقسم السلسلة الى نصفين ثم نحسب متوسط كل نصف وبذلك

نحصل على قيمتين نحدددهما على الرسم الخاص بالسلسلة ثم نصل بينهما بخط مستقيم نحسب ميله لاييجاد قيمة الاتجاه العام . ومن الواضح ان هذه الطريقة لا يمكن اتباعها الا اذا كان الاتجاه العام مستقيما .

ونلاحظ أنه اذا أخذنا متوسطات متحركة على أساس فترة زوجية (اربع سنوات مثلا) فان القيمة المتوسطة لا تكون سنة معينة وهذا ما يجعل التمهيدي قليل الفائدة حيث يكون هدفنا هو الحصول على القيم الاتجاهية (أي قيم الظاهرة وهي تتأثر فقط بالاتجاه العام أي قيمتها بعد استبعادات التغيرات الاخرى) المقابلة لكل سنة لذلك يمكن أخذ متوسط كل متوسطين الأول والثاني ، ثم الثاني والثالث وهكذا وبذلك نحصل على متوسطات تقابل سنوات معينة

مثال ٦٩ :

السنوات	القيم	متوسط متحرك لكل خمس سنوات
١٩٤٨	٥٠ر٠	
١٩٤٩	٣٦ر٥	
١٩٥٠	٤٣ر٠	٤٢ر٦
١٩٥١	٤٤ر٥	٤٠ر٢
١٩٥٢	٣٨ر٩	٣٩ر٤
١٩٥٣	٣٨ر١	٣٩ر٦
١٩٥٤	٣٢ر٦	٣٨ر٠
١٩٥٥	٣٨ر٧	٢٨ر٤
١٩٥٦	٤١ر٧	٣٧ر٦
١٩٥٧	٤١ر١	
١٩٥٨	٣٨ر٨	

السنوات	القيم	متوسط متحرك لكل اربعة سنوات	متوسط متحرك لكل متوسطين
١٩٤٨	٥٠٠٠		
١٩٤٩	٣٦٠٥	٤٣٠٥	٤٢٠١
١٩٥٠	٤٣٠٠		
١٩٥١	٤٤٠٥	٤٠٠٧	٤٠٠٩
١٩٥٢	٣٨٠٩	٤١٠١	٣٩٠٨
١٩٥٣	٣٨٠١	٣٨٠٥	٣٧٠٨
١٩٥٤	٣٢٠٦	٣٧٠١	٣٨٠٥
١٩٥٥	٣٨٠٧	٣٧٠٨	٣٨٠٢
١٩٥٦	٤١٠٧	٣٨٠٥	٣٨٠٧
١٩٥٧	٤١٠١	٣٨٠٨	
١٩٥٨	٣٣٠٨		

وبرسم السلسلة من واقع المتوسطات الأخيرة سواء على أساس خمس سنوات أو على أساس أربع سنوات نحصل على خط بياني ممد تقريباً وبذلك يكون من السهل تحديد اتجاهه العام وحساب ميله .

مثال ٧٠ :

إيجاد الاتجاه العام بتوفيق المستقيم بطريقة المربعات الصغرى :

السنوات	القيم (ص)	س	س ^٢	س ص
١٩٤٨	٥٠ر٠	صفر	صفر	صفر
١٩٤٩	٣٦ر٥	١	١	٣٦ر٥
١٩٥٠	٤٣ر٠	٢	٤	٨٦ر٠
١٩٥١	٤٤ر٥	٣	٩	١٣٣ر٥
١٩٥٢	٣٨ر٩	٤	١٦	١٥٥ر٦
١٩٥٣	٣٨ر١	٥	٢٥	١٩٠ر٥
١٩٥٤	٣٢ر٦	٦	٣٦	١٩٥ر٦
١٩٥٥	٣٨ر٧	٧	٤٩	٢٧٠ر٩
١٩٥٦	٤١ر٧	٨	٦٤	٣٣٣ر٦
١٩٥٧	٤١ر١	٩	٨١	٣٦٩ر٩
١٩٥٨	٣٨ر٨	١٠	١٠٠	٣٣٨ر٠
المجموع	٤٣٨ر٩	٥٥	٣٨٥	٢١١٠ر١

محد ص = م محد س + ن ب (١) $٥٥ = ٤٣٨ر٩ م + ١١ ب$

محد ص = م محد س^٢ + ب محد س (٢) $٥١١٠ر١ = ٣٨٥ م + ٥٥ ب$

بضرب المعادلة (١) في ٥ $٢١٩٤ر٥ = ٢٧٥ م + ٥٥ ب$

$٢١١٠ر١ = ٣٨٥ م + ٥٥ ب$

$٨٤ر٤ = ١١٠ م$

بالطرح

$\therefore م = \frac{٨٤ر٤}{١١٠} = ٠.٧٦٧$

أي ان الاتجاه العام لهذه السلسلة هبوطي ويساوي ٧٦٧ر٠ في العام الواحد .

بالتعويض في المعادلة (١) $٤٣٨٩ = - ٤٢١٨٥ + ١١ ب$

$$٤٨١٠٨٥ = ١١ ب$$

$$٤٣٧٣٥ = ب$$

∴ معادلة الاتجاه العام من نقطة الاصل عام ١٩٤٨

$$ص = ٧٦٧ر٠ س + ٤٣٧٣٥$$

مثال ٧١ :

ويمكن أخذ عام ١٩٥٣ كنقطة أصل وبذلك يكون منه ل كالاتي :

السنوات	القيم (ص)	س	س ص	س
١٩٤٨	٥٠ر٠	٥ -	٢٥٠ -	٢٥
١٩٤٩	٣٦ر٥	٤ -	١٤٦ -	١٦
١٩٥٠	٤٣ر٠	٣ -	١٢٩ -	٩
١٩٥١	٤٤ر٥	٢ -	٨٩ -	٤
١٩٥٢	٣٨ر٩	١ -	٣٨٩ -	١
١٩٥٣	٣٨ر١	صفر	صفر	صفر
١٩٥٤	٢٢ر٦	١ +	٣٢٦ +	١
١٩٥٥	٣٨ر٧	٢ +	٧٧٤ +	٤
١٩٥٦	٤١ر٧	٣ +	١٢٥١ +	٩
١٩٥٧	٤١ر١	٤ +	١٦٤ر٤ +	١٦
١٩٥٨	٣٣ر٨	٥ +	١٦٩ر٠ +	٢٥
المجموع	٤٣٨ر٩	صفر	٧٤ر٤ -	١١٠

$$\therefore \text{ب} = ٣٩٩$$

$$٤٣٨٩ = ١١ \text{ ب}$$

$$\therefore \text{م} = - ٠٧٦٧$$

$$- ٨٤٤ = ١١٠ \text{ م}$$

∴ معادلة الاتجاه العام من نقطة الاصل عام ١٩٥٣

$$\text{ص} = - ٠٧٦٨ \text{ س} + ٣٩٩$$

(نلاحظ أن قيمة م لم تتغير حيث ان ميل خط الاتجاه العام لا يمكن أن يتغير ، أما قيمة ب فتغيرت حيث انها تمثل الجزء المقطوع من المحور الرأسي بواسطة خط الاتجاه العام وهذا الجزء لا بد ان يتغير إذا غيرنا نقطة الأصل الخاصة بالسلسلة) .

وعلى أساس معادلة الاتجاه العام يمكن حساب القيم الاتجاهية أي قيم الظاهرة وهي متأثرة فقط بالاتجاه العام وذلك بالتعويض في معادلة الاتجاه عن س حسب ترتيب السنوات .

فإذا أردنا حساب القيمة الاتجاهية لعام ١٩٤٨ يكون العمل كالآتي :

$$\text{ص} = - ٠٧٦٧ \times - ٥ + ٣٩٩$$

$$= ٣٨٣٥ + ٣٩٩$$

$$= ٤٣٧٣٥$$

وسواء أخذنا هذه المعادلة أو المعادلة السابقة فاننا نحصل على نفس النتيجة للقيم الاتجاهية :

$$\text{ص} = - ٠٧٦٧ \times \text{صفر} + ٤٣٧٣٥$$

$$= ٤٣٧٣٥$$

وتكون القيم الاتجاهية لجميع السنوات كالآتي : ٤٣٧ ، ٤٣٠ ،

٤٣٢ ، ٤١٤ ، ٤٠٧ ، ٣٩٩ ، ٣٩١ ، ٣٨٤ ، ٣٧٦ ، ٣٦٨ ،

٣٦١ .

وإذا كان عدد السنوات زوجياً فلا نستطيع أن نختصر العمل كثيراً
 بأخذ نقطة الأصل في الوسط حيث لا توجد سنة متوسطة يمكن أن تعتبر
 نقطة الأصل ، ولذلك تؤخذ نقطة الأصل في منتصف الفئة بين السنتين
 المتوسطتين وبذلك يكون وحدات الزمن هي نصف السنة ويتضح ذلك من
 المثال الآتي :

مثال ٧٢ :

السنوات	القيم (ص)	س	س ^٢	س ص
١٩٤٨	٥٠ر٠	٩ -	٨١	٤٥٠ر٠ -
١٩٤٩	٣٦ر٥	٧ -	٤٩	٢٥٥ر٥ -
١٩٥٠	٤٣ر٠	٥ -	٢٥	٢١٥ر٠ -
١٩٥١	٤٤ر٥	٣ -	٩	١٣٣ر٥ -
١٩٥٢	٣٨ر٩	١ -	١	٣٨ر٩ -
		صفر	صفر	صفر
١٩٥٣	٣٨ر١	١	١	٣٨ر١
١٩٥٤	٣٢ر٦	٣	٩	٩٧ر٨
١٩٥٥	٣٨ر٧	٥	٢٥	١٩٣ر٥
١٩٥٦	٤١ر٧	٧	٤٩	٢٩١ر٩
١٩٥٧	٤١ر١	٩	٨١	٣٦٩ر٩
المجموع	٤٠٥ر١	صفر	٣٣٠	١٠١٧ -

$$\therefore \text{ب} = ٤٠ر٥١$$

$$\therefore \text{م} = ٠٣٠٨$$

$$\therefore ١٠ = ٤٠٥ر١$$

$$- ١٠١٧ = ٣٣٠ \text{ م}$$

وبذلك يكون الاتجاه العام النصف سنوي مبطوي ويساوي ٠.٣٠٨ في المتوسط فاذا أردنا حساب القيمة الاتجاهية لعام ١٩٤٨ يكون العمل كالآتي : -

$$ص = ٣٠٨ \times ٩ - ٤٠٥١$$

$$= ٢٧٧٢ + ٤٠٥١$$

$$= ٤٣٢٨٢$$

وعلى ساس معادلة الاتجاه العام يمكن للتنبؤ بما ستكون عليه قيمة الظاهرة في المستقبل على فرض ثبات العوامل المختلفة المؤثرة على هذه الظاهرة فاذا أردنا تقدير كم ستكون عليه قيمة الظاهرة عام ١٩٥٨ واستخدمنا المعادلة الأخيرة التي تكون فيها س ممثلة لوحداث نصف سنوية يكون العمل كالآتي :

$$ص = ٣٠٨ \times ١١ - ٤٠٥١$$

$$= ٣٣٨٨ + ٤٠٥١$$

$$= ٣٧١٣٢$$

ولا شك أن توفيق خط الاتجاه العام بطريقة المربعات الصغرى ادى بكثير من طريقة المتوسطات المتحركة . على أن دقة نتائجها تتوقف على تحديد نوع الخط الذي يناسب تغير الظاهرة في المدة موضوع البحث ، وبشكل آخر تتوقف على تحديد المعادلة التي تصاح لتمثيل الاتجاه العام للظاهرة .

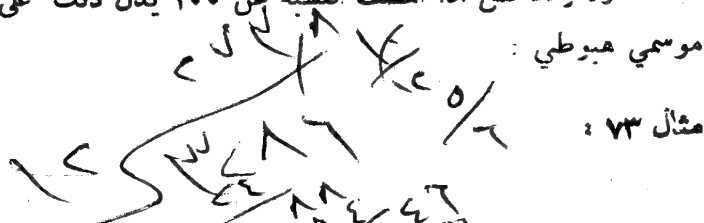
حساب التغيرات الموسمية :

حيث أن التغيرات الموسمية تحدث في أجزاء السنة ، لذلك يجب أن تكون قيم الظاهرة على أساس الأجزاء التي نريد حساب التغيرات الموسمية الخاصة بها ، فاذا كما نريد حساب التغيرات الموسمية في أشهر السنة (مثلا)

يجب أن تكون قيم الظاهرة متوفرة على أساس شهري ، وإذ كنا نريد حساب التغيرات الموسمية الربع سنوية يجب أن تكون القيم ربع سنوية...الخ.
وهناك عدة طرق لحساب التغيرات الموسمية غير أننا سنكتفي بشرح طريقتين منها .

١ - إيجاد النسب الموسمية بطريقة المتوسطات البسيطة .

تعتمد هذه الطريقة على حساب المتوسط العام لكل فترة خلال المدة موضوع البحث ثم حساب نسبة قيمة الظاهرة في كل فترة الى هذا المتوسط العام فإذا زادت النسبة عن ١٠٠ يدل ذلك على وجود تغير موسمي صعودي في هذه الفترة وإذا نقصت النسبة عن ١٠٠ يدل ذلك على وجود تغير موسمي هبوطي :



السنوات يناير فبراير مارس أبريل مايو يونيو يوليو أغسطس سبتمبر أكتوبر نوفمبر ديسمبر

١٩٥١	٣١٨	٢٨١	٢٧٨	٢٥٠	٢٣١	٢١٦	٢٢٣	٢٤٥	٢٦٩	٣٠٢	٣٢٥	٣٤٧
١٩٥٢	٣٤٢	٣٠٩	٢٩٩	٢٦٨	٢٤٩	٢٣٦	٢٤٢	٢٦٢	٢٨٨	٣٢١	٣٤٢	٣٦٤
١٩٥٣	٣٦٧	٣٢٨	٣٢٠	٢٨٧	٢٦٩	٢٥١	٢٥٩	٢٨٤	٣٠٩	٣٤٥	٣٦٧	٣٩٤
١٩٥٤	٣٩٢	٣٤٩	٣٤٢	٣١١	٢٩٠	٢٧٣	٢٨٢	٣٠٥	٣٢٨	٣٦٤	٣٨٩	٤١٧
١٩٥٥	٤٢٠	٣٧٨	٣٧٠	٣٣٤	١١٤	٢٩٦	٣٠٥	٣٣٠	٣٥٦	٣٩٦	٤٤٢	٤٥٢
١٩٥٦	٤٥٣	٤١٢	٣٩٨	٣٦٢	٣٤١	٣٢٢	٢٢٥	٣٥٩	٣٩٢	٤٢٧	٤٥٤	٤٨٣
١٩٥٧	٣٨٧	٤٤٠	٤٢٩	٣٩٣	٣٧٠	٣٤٧	٣٥٧	٣٨٨	٤١٥	٤٥٧	٤٩١	٥١٦
١٩٥٨	٥٢٩	٤٧٧	٤٦٣	٤٢٣	٣٩٨	٣٨٠	٣٨٩	٤١٩	٤٤٨	٤٩٣	٥٢٦	٥٦٠

المجموع الشهري :

٣١٠٥ ٢٨٠٥ ٢٥٩٢ ٢٣٨٢ ٢٣٢١ ٢٤٦٢ ٢٦٢٨ ٢٨٩٩ ٢٩٧٤ ٣٣٠٨
٣٥٣٣ ٣٣١٦

المتوسط الشهري

٤٤١ ٤١٤ ٣٨٨ ٣٥٠ ٣٢٤ ٢٩٩ ٢٩٠ ٣٠٧ ٣٢٨ ٣٦٢ ٣٧٢ ٤١٣

وحيث ان المتوسط الشهري العام = ٣٥٧ .

∴ النسب الموسمية الشهرية هي :

$$\text{أي صعود موسمي بنسبة } ١٦\% \quad ١١٦ = ١٠٠ \times \frac{٤١٣}{٣٥٧} \text{ تقريباً}$$

$$\text{أي صعود موسمي بنسبة } ٤\% \quad ١٠٤ = ١٠٠ \times \frac{٣٧٢}{٣٥٧} \text{ تقريباً}$$

$$\text{أي صعود موسمي بنسبة } ١\% \quad ١٠١ = ١٠٠ \times \frac{٣٦٢}{٣٥٧} \text{ تقريباً}$$

$$\text{أي هبوط موسمي بنسبة } ٩\% \quad ٩١ = ١٠٠ \times \frac{٣٢٨}{٣٥٧} \text{ تقريباً}$$

وهكذا بالنسبة للأشهر الباقية .

وإذا كانت قيم الظاهرة متأثرة باتجاه عام معين فلا بد من استبعاد هذا الاتجاه قبل حساب النسب الموسمية حيث اننا إذا لم نستبعده فإن التغيرات الموسمية لا تكون دقيقة ، فإذا كان الاتجاه العام صعودياً ، وكان التغير الموسمي صعودياً كذلك فإن النسب الموسمية تكون أكثر مما يجب أن تكون والعكس إذا كان الاتجاه العام صعودياً وكان التغير الموسمي هبوطياً فإن

النسب الموسمية تكون أقل مما يجب أن تكون ، ولاستبعاد الاتجاه العام نوجد معادلة الاتجاه العام للقيم الموجودة لدينا ومنها نحسب القيم الاتجاهية للظاهرة ثم نقسم كل قيمة واقعية على القيمة الاتجاهية المقابلة لها (نسبة مئوية) فنحصل على نسب تدل على الظاهرة مستبعداً منها الاتجاه العام .

مثال ٧٤ :

(من التمرين السابق)

السن	س	س	المتوسط الشهري	السنوات
٤٩	١٩١٥٠٩ -	٧ -	٢٧٣٠٧	١٩٥١ ✓
٢٥	١٤٦٧٠٥ -	٥ -	٢٩٣٠٥	١٩٥٢ ✓
٩	٩٤٥٠٠ -	٣ -	٣١٥٠٠	١٩٥٣ ✓
١	٣٣٦٠٨ -	١ -	٣٣٦٠٨	١٩٥٤ ✓
١	١٦٤٠٤	١	٢٦٤٠٤	١٩٥٥ ✓
٩	١١٨٤٠٤	٣	٣٩٤٠٨	١٩٥٦ ✓
٢٥	٢١٢١٠٠	٥	٤٢٤٠٢	١٩٥٧ ✓
٤٩	٣٢١٠٠٩	٧	٤٥٨٠٧	١٩٥٨ ✓
١٦٨	٢٢١٥٠٥	صفر	٢٨٦١٠١	المجموع

$$\frac{\text{مجموع القيم في العام}}{١٢} = \text{المتوسط الشهري}$$

$$\text{م ح ص} = \text{م ح س} + \text{ن ب}$$

$$\text{م ح س ص} = \text{م ح س} + \text{ن ب م ح س}$$

$$\frac{28611}{3576} = \text{ب} = \frac{28611}{8}$$

$$\frac{22155}{13188} = \text{م} = \frac{22155}{168}$$

∴ معادلة الاتجاه العام = 13188 م + 3576 ب
حيث م تمثل وحدات نصف سنوية من نقطة الأصل 1 يناير 1900 ،

$$\text{وبذلك يكون الاتجاه العام الشهري} = \frac{13188}{12} = 22 \text{ تقريباً} .$$

وعلى أساس أن القيم الشهرية في التعرین السابق هي على أساس منتصف الشهر تكون القيمة الاتجاهية لمنتصف شهر يناير 1901 =

$$3576 + 475 - \times 22 =$$

$$3576 + 1045 =$$

$$2531 =$$

$$\text{والقيمة الاتجاهية لمنتصف شهر فبراير 1901} = 3576 + 475 - \times 22 =$$

$$2503 =$$

$$\text{والقيمة الاتجاهية لمنتصف شهر مارس 1901} = 3576 + 450 - \times 22 =$$

$$2575 =$$

$$\text{والقيمة الاتجاهية لمنتصف شهر يناير 1900} = 3576 + 0 - \times 22 =$$

$$3587 =$$

$$\text{والقيمة الاتجاهية لمنتصف شهر فبراير 1900} = 3576 + 150 \times 22 =$$

$$3609$$

وهكذا بالنسبة للشهر الباقية) نلاحظ ان نقطة الأصل هي 1 يناير 1900

وبذلك يكون ترتيب منتصف شهر يناير ١٩٥١ = ٤٧٥، وترتيب منتصف شهر فبراير ١٩٥١ = ٤٦٥ وهكذا ...) . وتكون بذلك القيم الاتجاهية لجميع الأشهر موضوع البحث كالآتي :

السنوات	يناير	فبراير	مارس	ابريل	مايو	يونيو	يوليو	اغسطس	سبتمبر	اكتوبر	نوفمبر	ديسمبر
١٩٥١	٢٥٣	٢٥٥	٢٥٧	٢٦٠	٢٦٢	٢٦٤	٢٦٦	٢٦٨	٢٧٠	٢٧٣	٢٧٥	٢٧٧
١٩٥٢	٢٧٩	٢٨٢	٢٨٤	٢٨٦	٢٨٨	٢٩٠	٢٩٣	٢٩٥	٢٩٧	٢٩٩	٣٠١	٣٠٣
١٩٥٣	٣٠٦	٣٠٨	٣١٠	٣١٢	٣١٥	٣١٧	٣١٩	٣٢١	٣٢٣	٣٢٦	٣٢٨	٣٣٠
١٩٥٤	٢٣٢	٢٣٤	٢٣٧	٢٣٩	٢٤١	٢٤٣	٢٤٥	٢٤٧	٢٥٠	٢٥٢	٢٥٤	٢٥٦
١٩٥٥	٣٥٩	٣٦١	٣٦٣	٣٦٥	٣٦٧	٣٧٠	٣٧٢	٣٧٤	٣٧٦	٣٧٨	٣٨١	٣٨٣
١٩٥٦	٣٨٥	٣٨٧	٣٨٩	٣٩٢	٣٩٤	٣٩٦	٣٩٨	٤٠٠	٤٠٣	٤٠٥	٤٠٧	٤٠٩
١٩٥٧	٤١١	٤١٤	٤١٦	٤١٨	٤٢٠	٤٢٢	٤٢٥	٤٢٧	٤٢٩	٤٣١	٤٣٣	٤٣٦
١٩٥٨	٤٣٨	٤٤٠	٤٤٢	٤٤٤	٤٤٧	٤٤٩	٤٥١	٤٥٣	٤٥٥	٤٥٨	٤٦٠	٤٦٢

وبقسمة القيم الشهرية الأصلية على القيم الاتجاهية نحصل على النسب الآتية :

السنوات	يناير	فبراير	مارس	ابريل	مايو	يونيو	يوليو	اغسطس	سبتمبر	اكتوبر	نوفمبر	ديسمبر
١٩٥١	١٢٦	١١٠	١٠٨	٩٦	٨٨	٨٢	٨٤	٩١	٩٩	١١١	١١٨	١٢٥
١٩٥٢	١٢٢	١١٠	١٠٥	٩٤	٨٦	٨١	٨٣	٨٩	٩٧	١٠٧	١١٣	١٢٠
١٩٥٣	١٢٠	١٠٦	١٠٣	٩٢	٨٥	٧٩	٨١	٨٨	٩٥	١٠٦	١١٢	١١٩
١٩٥٤	١١٨	١٠٤	١٠٢	٩٢	٨٥	٧٩	٨٢	٨٨	٩٤	١٠٣	١١٠	١١٧
١٩٥٥	١١٧	١٠٥	١٠٢	٩١	٨٥	٨٠	٨٢	٨٨	٩٥	١٠٥	١١١	١١٨
١٩٥٦	١١٨	١٠٦	١٠٢	٩٢	٨٧	٨١	٨٤	٩٠	٩٧	١٠٥	١١٢	١١٨
١٩٥٧	١١٨	١٠٦	١٠٣	٩٤	٨٨	٨٢	٨٤	٩١	٩٧	١٠٦	١١٣	١١٨
١٩٥٨	١٢١	١٠٨	١٠٥	٩٥	٨٩	٨٥	٨٦	٩٢	٩٨	١٠٨	١١٤	١٢١

وبذلك تكون المتوسطات الشهرية هي كالآتي :

١٢٠ ، ١٠٧ ، ١٠٤ ، ٩٣ ، ٨٧ ، ٨١ ، ٨٤ ، ٩٠ ، ٩٦ ، ١٠٦ ،

١١٣ ، ١١٩ ، وحيث ان مجموع هذه النسب = ١٢٠٠ فتكون النسب السابقة هي نفسها تعبر عن النسب الموسمية ، أما إذا كان مجموعها ليس ١٢٠٠ فيجب تصحيحها حتى نحصل على النسب الموسمية ، فإذا فرضنا أن مجموعها

$$= ١١٠٠ ، فإن النسبة الاولى تصبح بعد التصحيح ١٢٠ × \frac{١٢٠٠}{١١٠٠} = ١٣١$$

تقريباً ، وهكذا بالنسبة لباقي المتوسطات الشهرية .

٢ - إيجاد التقلبات الموسمية بوحدات مطلقة .

قد نحتاج أحياناً إلى حساب القيم الفعلية للتقلبات الموسمية . ولحساب هذه القيم نتبع الخطوات الآتية :

١ - نحسب القيم الاتجاهية سواء على أساس المتوسطات المنحركة أو على أساس معادلة الاتجاه العام .

٢ - نحسب انحرافات القيم الفعلية عن القيم الاتجاهية فيكون الفرق هو عبارة عن التغيرات الموسمية والعرضية والدورية ، فإذا كنا قد أخذنا فترة ليس فيها تغيرات عرضية أو دورية يكون الفرق هو فقط التغيرات الموسمية ، أما إذا كان هناك تغيرات عرضية نحسب متوسط الانحرافات ، وبذلك نقضي على هذه التغيرات حيث أن التقلبات العرضية لا تتبع قاعدة معينة فأحياناً تكون صعودية وأحياناً تكون هبوطية ، وبذلك يتلأشى أثرها عند حساب المتوسط (هذا على أساس أن نكون قد أخذنا مدة طويلة بحيث يمكن أن تكون هناك تغيرات عرضية صعودية وأخرى هبوطية) .

مشال ٧٥ :

انحراف القيمة الواقعة عن القيمة الاتجاهية	متوسط متحرك على اساس فترتين	متوسط متحرك على اساس اربع سنوات	القيم	فترات الربع سنوية
			٦٥	١٩٥١ الربع الاول
			٦٢	د الثاني
١٥٠ -	٦٢٥	٦٣	٦١	د الثالث
١٥٠ +	٦١٥	٦٢	٦٣	د الرابع
٣٥٠ +	٦٠٥	٦١	٦٤	١٩٥٢ د الاول
٢٥٠ -	٥٩٥	٦٠	٥٧	د الثاني
٣٥٠ -	٥٩٥	٥٩	٥٦	د الثالث
٥٥٠ +	٦٠٥	٦٠	٦١	د الرابع
٦٠٠ +	٦٢	٦١	٦٨	١٩٥٣ د الاول
١٥٠ -	٦٣٥	٦٣	٦٢	د الثاني
٣٥٠ -	٦٤٥	٦٤	٦١	د الثالث
٢٥٠ +	٦٤٥	٦٥	٦٧	د الرابع
٦٥٠ +	٦٣٥	٦٤	٧٠	١٩٥٤ د الاول
٣٥٠ -	٦٢	٦٣	٥٩	د الثاني
٤٥٠ -	٦٠	٦١	٥٦	د الثالث
٣٥٠ +	٥٧	٥٩	٦١	د الرابع
٢٥٠ +	٥٦٥	٥٧	٥٩	١٩٥٥ د الاول
١٥٠ -	٥٥٥	٥٦	٥٤	د الثاني
		٥٥	٥١	د الثالث
			٥٨	د الرابع

فاذا اخذنا الربع الاول من كل عام وحسبنا متوسط الانحرافات يكون الناتج هو التغير الموسمي ، $35 + 6 + 65 + 25 + 25$ ومتوسطها 46 وهو قيمة التغير الموسمي في هذا الربع - وفي الربع الثاني - $25 - 15 - 3 - 3 = 15$ ومتوسطها 21 وهو قيمة التغير الموسمي في هذا الربع . وفي الربع الثالث - $15 - 35 - 35 - 4$ ومتوسطها 2875 وهو قيمة التغير الموسمي في هذا الربع . وفي الربع الرابع $15 - 35 - 35 + 1875$ وهو قيمة التغير الموسمي في هذا الربع .

استبعاد تأثير التقلبات الموسمية :

من الواضح ان الهدف من دراسة التقلبات الموسمية هو تحديد تأثير كل موسم على قيم الظاهرة موضوع البحث حتى يمكن الاستعداد لمواجهة الزيادة الموسمية او النقص الموسمي ولهذا اهمية كبيرة بالنسبة للمنتجين او التجار الذين تتعرض منتجاتهم لتقلبات موسمية . ومن ناحية اخرى فان المنتج يهتم كذلك ان يتعرف على الحالة الحقيقية للطلب على منتجاته ، اي حالة الطلب بدون التأثيرات الموسمية ، المختلفة حتى لا تضلله الزيادة المؤقتة او النقص المؤقت في مبيعاته ، ويمكن استبعاد تأثير الموسم بقسمة القيمة الواقعية للظاهرة في الفترة موضوع البحث على النسبة الموسمية في هذه الفترة . فاذا فرضنا مثلا ان قيمة مبيعات محل تجاري في شهر معين هي ٥٠٠٠٠ ليرة وكانت النسبة الموسمية للمبيعات في هذا الشهر هي ١٢٥٪ تكون مبيعات المحل بدون تأثير الموسم

أي بعد استبعاده $= \frac{100}{125} \times 50000 = 40000$ ليرة . أما اذا كانت النسبة

الموسمية تدل على هبوط مثلا ٨٠٪ تكون قيمة المبيعات بدون تأثير

الموسم $= \frac{100}{80} \times 50000 = 62500$ واذا كانت التغيرات الموسمية معروفة

في شكل قيم مطلقة يكون استبعادها بطرحها مباشرة من القيم الحقيقية للظاهرة (طرحا جبريا) .

التغيرات العرضية والدورية :

إذا توفرت لدينا القيم الواقعية والاتجاهية لظاهرة ما في فترة زمنية معينة فإن الفرق بينها يدلنا على التغيرات الموسمية والعرضية ، فإذا طرحنا قيمة التغير الموسمي يكون الباقي هو التغير العرضي في هذه الفترة .

مثال ٧٦ :

إذا كانت القيمة الواقعية = ٦٤ والقيمة الاتجاهية = ٦٥ والتغير الموسمي = - ٣

$$\text{التغير العرضي} = \frac{(65 - 3) + 64}{2} - 64 = \frac{62}{2} - 64 = 31 - 64 = -33$$

وإذا كانت القيمة الواقعية = ٦٦ والقيمة الاتجاهية = ٦٢ والتغير الموسمي = + ٤

$$\therefore \text{التغير العرضي} = 66 - (62 + 4) = 66 - 66 = 0$$

$$66 - 66 = 0$$

وإذا لم تكن الظاهرة قد أصيبت بتغيرات عرضية يكون الفرق بين القيمة الواقعية والقيمة الاتجاهية هو التغيرات الموسمية والتغيرات الدورية ، فإذا طرحنا قيمة التغير الموسمي يكون الباقي هو التغير الدوري . كذلك يمكن بقسمة القيمة الواقعية للظاهرة على قيمتها الاتجاهية أن نحصل على نسبة مئوية تدل على الظاهرة بدون الاتجاه العام ، وإذا قسمنا هذه النسبة على النسبة

الموسمية في الفترة موضوع البحث نحصل على نسبة مئوية تدل على التغير الدوري والعرضي سوياً ، فإذا كانت لدينا فترة طويلة يمكننا أن نستبعد التغيرات العرضية بحساب المتوسطات المتحركة للنسب التي تحصل عليها فتكون النسب بعد استبعاد التغيرات العرضية معبرة عن التغيرات الدورية . وبذلك نستطيع أن نستنتج القاعدة بأن القيمة المشاهدة (الواقعية) = القيمة الاتجاهية \times التغير الموسمي \times التغير الدوري .

مثال ٧٠ :

الأسهر	القيم الواقعية	القيم الاتجاهية	القيمة الواقعية الاتجاهية $\times 100$	النسب الموسمية	النسب الدورية
يناير	١٦٩	٣٩٥	٤٣ %	٧٥ %	٥٧ %
فبراير	٢١٢	٣٩٦	٥٤	٩٦	٥٦
مارس	٢٥١	٣٩٧	٦٣	١٠٧	٥٩
ابريل	٢٨٦	٣٩٨	٧٢	١٢٣	٥٨
مايو	٢٨٥	٣٩٩	٧١	١٢٢	٥٨
يونيو	٢٦٤	٤٠٠	٦٦	١١١	٥٩
يوليو	١٦٤	٤٠١	٤٩	٨٣	٤٩
اغسطس	١٩٦	٤٠٢	٤٩	٧٢	٦٨
سبتمبر	٢٠٨	٤٠٣	٥٢	٨٨	٥٩
اكتوبر	٢٦٤	٤٠٤	٦٥	١٠٨	٦٠
نوفمبر	٢٧٤	٤٠٥	٦٨	١١٠	٦٢
ديسمبر	٢٧٣	٤٠٦	٦٧	١٠٣	٦٥

نسب التغيرات الدورية = النسبة بدون الاتجاه العام ÷ النسبة الموسمية .
وتطبيقاً للقاعدة التي سبق الإشارة إليها - القيمة المشاهدة = القيمة
الاتجاهية × التغير الموسمي × التغير الدوري × التغير العرضي يمكن التنبؤ بما
سوف تكون قيمة الظاهرة في فترة مستقبلية . ولما كان التغير العرضي لا يمكن
التنبؤ به لهذا يكون تنبؤنا قاصراً على قيمة الظاهرة وهي متأثرة فقط بالاتجاه
العام والتغير الموسمي والتغير الدوري .

مثال ٧٨ :

إذا كانت القيمة الاتجاهية المتوقعة في شهر ما هي ٣٩٥ وان النسبة
الموسمية لهذا الشهر هي ٧٥٪ والنسبة الدورية ٥٧٪ .

∴ قيمة الظاهرة وهي متأثرة بهذه الأنواع الثلاث من التغيرات .

$$= 395 \times \frac{75}{100} \times \frac{57}{100} = 168.86 \text{ أي } 169 \text{ تقريباً .}$$

الارتباط بين السلاسل الزمنية :

تبين لنا مما تقدم ان السلاسل الزمنية تتأثر بأربع أنواع من التغيرات -
الاتجاه العام والتقلبات الموسمية والدورية والعرضية . وعندما نريد دراسة
الارتباط بين سلسلتين زمنيتين يجب أن نحدد نوع العلاقة التي نريد دراستها ،
فقد ترغب في بحث ما إذا كان هناك ارتباط بين الاتجاه العام لسلسلة ما
والاتجاه العام لسلسلة أخرى . وإذا كان هذا هو هدفنا نحسب القيم الاتجاهية
لكل من السلسلتين ثم نرصد هذه القيم على ورق الرسم البياني وسوف يتضح
من الرسم مباشرة إذا كان هناك اتفاق بين اتجاهي الظاهرتين ، الأمر الذي يدل
على وجود الارتباط بينهما ، وفي هذه الحالة لا يكون هناك ضرورة لحساب
معامل الارتباط حيث أن مجرد الاتفاق بين الاتجاه العام للظاهرتين يدل على

وجود ارتباط قوي مها كانت درجة المخدار كل من خطي الاتجاه العام .

على أن الذي يهنا عند دراسة الارتباط بين سلسلتين هو الاجابة على السؤال - هل هناك علاقة بين ذبذبة سلسلة ما حول اتجاهها العام وذبذبة سلسلة أخرى حول اتجاهها العام أو لا ، وما درجة هذه العلاقة ؟

فاذا كان هذا هو هدفنا يتحتم علينا حسب قيم هذه الذبذبات في كل من السلسلتين ثم نقوم بحساب الارتباط بين هذه القيم باي قانون من قوانين الارتباط .

مثال ٧٩ :

الرقم القياسي لأسعار الأسهم الصناعية	الرقم القياسي للاتنتاج الصناعي	السنوات
١٠٩	١٠١	١٩٥٤
٨٢	٦٧	١٩٥٥
١٢٠	٨٥	١٩٥٦
١٠٣	١٠٥	١٩٥٧
٩٠	٩٨	١٩٥٨
٩٦	١١٠	١٩٥٩
٩٩	١٠٨	١٩٦٠

١ - نحسب اولا القيم الاتجاهية لكل من السلسلتين بطريقة المربعات الصغرى .

معادلة الاتجاه العام للرقم القياسي للإنتاج الصناعي :

س	ص	س ص	س ^٢
٣ -	١٠١	٣٠٣ -	٩
٢ -	٦٧	١٣٤ -	٤
١ -	٨٥	٨٥ -	١
صفر	١٠٥	صفر	صفر
١	٩٨	٩٨	١
٢	١١٠	٢٢٠	٤
٣	١٠٨	١٢٤	٩
المجموع	٦٧٤	١٢٠ +	٢٨

$$\begin{aligned} \text{مح ص} &= \text{م مح س} + \text{ن} \\ \text{مح ص} &= \text{م مح س} + \text{ب مح س} + \text{م} \end{aligned}$$

$$\therefore \text{ب} = \frac{674}{7} = 962857$$

$$\text{و م} = \frac{120}{28} = 42857$$

$$\therefore \text{ص} = 962857 + 42857 \text{ س}$$

معادلة الاتجاه العام للرقم القياسي لأسعار الاسهم الصناعية

س	ص	ص س	س ^٢
٣ -	١٠٩	٣٢٧ -	٩
٢ -	٨٢	١٦٤ -	٤
١ -	١٢٠	١٢٠ -	١
صفر	١٠٣ .	صفر	صفر
١	٩٠	٩٠	١
٢	٩٦	١٩٢	٤
٣	٩٩	٢٩٧	٩
المجموع	٦٩٩	٣٢ -	٢٨

$$\text{مح ص} = \text{مح م س} + \text{ن ب} \quad \text{ب } ٧ = ٦٩٩$$

$$\text{مح س ص} = \text{مح م س}^٢ + \text{ب مح س} \quad \text{م } ٢٨ = ٣٢ -$$

$$٩٩٨٥٧ = \frac{٦٩٩}{٧} = \text{ب} \therefore$$

$$١١٤٣ - = \frac{٣٢}{٢٨} = \text{م} \text{ و}$$

$$\text{ب} \therefore \text{ص} = ٩٩٨٥٧ + \text{س } ١١٤٣ -$$

٢ - نحسب القيم الاتجاهية بالتعويض في معادلات الاتجاه العام عن ص
ثم نحسب انحرافات القيم الواقعية عن القيم الاتجاهية :

انحرافات قيم الانتاج الصناعي	انحرافات قيم الاسهم	مربع انحرافات قيم الانتاج الصناعي	مربع انحرافات قيم الأسهم	حاصل ضرب الانحرافات
١٧٦ +	٥٧ +	٣٠٩٧٦	٣٢٤٩	١٠٠٣٢ +
١٠٧ -	٢٠١ -	٤٢٨٤٩	٤٠٤٠١	٤١٦٠٧ +
٧٠ -	١٩٠ +	٤٩٠٠	٣٦١٠٠	١٣٣٠٠ -
٨٧ +	٣٤ +	٧٥٦٩	٩٦١	٢٦٩٧ +
٢٦ -	٨٧ -	٦٧٦	٧٥٦٩	٢٢٦٢ +
٥١ +	١٦ -	٢٦٠١	٢٥٦	٨١٦ -
١١ -	٢٦ +	١٢١	٦٧٦	٢٨٦ -
المجموع	-	٨٩٦٩٢	٨٩٢١٢	٩٢١٩٦ +

(نلاحظ ان مجموع الانحرافات عن القيم الاتجاهية = صفر لكل من
الظاهرتين)

∴ الانحراف المعياري لانحرافات الرقم القياسي للانتاج الصناعي =

$$= \sqrt{\frac{٨٩٦٩٢}{١١٣}}$$

والانحراف المعياري لانحرافات الرقم القياسي لأسعار الأسهم =

$$112 = \frac{89212}{7}$$

$$\therefore R = \frac{42196}{88592} = \frac{42196}{112 \times 113 \times 7} = 0.47$$

تمارين

١ - الجدول الآتي يبين انتاج الصلب في احدى الدول :

السنة	الانتاج	السنة	الانتاج
١٩٥٠	٣٢١٠ طن	١٩٥٥	٥٨٣٠ طن
١٩٥١	٣٤٢٦ طن	١٩٥٦	٥٩٧٠ طن
١٩٥٢	٣٦٦٩ طن	١٩٥٧	٦٢٠٠ طن
١٩٥٣	٤٩٠٠ طن	١٩٥٨	٦٦٥٠ طن
١٩٥٤	٥٣٠٠ طن	١٩٥٩	٧٢٥٠ طن

اوجد معادلة الاتجاه العام بطريقة المربعات الصغرى ، ومنها استنتج القيم الاتجاهية في الأعوام ١٩٤٥ ، ١٩٥٥ ، ١٩٦٥ .

٢ - الأرقام الآتية تبين النسب النموذجية للتقد المتداول :

يناير	٩٥	يوليو	٩٣
فبراير	٩٩	اغسطس	٩٨
مارس	١٠٠	سبتمبر	١٢٠

ابريل	٩٩	اكتوبر	١٠٩
مايو	٩٤	نوفمبر	٩٧
يونيو	٩٢	ديسمبر	٩٦

احسب النسب الموسمية .

٣ - احسب الارتباط بين الاتجاه العام للسلسلتين الآتيتين ، ثم بين تقلباتهما حول الاتجاه العام .

السنوات	الرقم القياسي للانتاج الزراعي	الرقم القياسي لاسعار الحاصلات الزراعية
١٩٥٠	١٣٦	٨٩
١٩٥١	٧٦	٦٤
١٩٥٢	١١٤	٦٥
١٩٥٣	١٣٤	٧٦
١٩٥٤	٨٣	٧٥
١٩٥٥	١٠٢	٩٢
١٩٥٦	٨٨	١١٨
١٩٥٧	٩٦	١١٥
١٩٥٨	١٢٠	١١٨

أعداد عشوائية

١٥	٤٩	٢٢	٠٢	٧٧	٩٦	٦٣	٤٨	٣٢	٩٨	٩٥	١٦	٥٣	٥٠	٣٢
٢٨	١٢	٣٦	٦٧	٦٤	٣٢	٤٠	٣٦	٤٠	٩٦	٨٢	٥١	٤٠	٥٢	٩٢
٣٤	٢٥	١١	٥٥	١٢	٥٠	٢٧	٤٣	٣٩	٠٣	٥٩	٣٤	٢١	٧٠	٢٧
٢٣	٨٢	٥٢	٣٧	٢٦	٥٤	٠٠	٧١	٥٣	٤٣	٩١	٧٠	١٦	١٠	٢٥
٦٧	٨٣	٨١	٤٢	٣٧	١٤	٤٩	٤٣	٠٦	٠١	٨٣	٤٩	٣٣	١١	٣١
٧٦	٤١	٢٦	١٧	٤٤	٢٥	١٢	٧٤	٢٥	٧٦	١٩	٧٨	٧٧	٥٤	٤٠
٨٥	٥٩	٤٥	٧٦	٦٤	٧٥	٢٠	٧٨	١٥	٣٤	١٧	٤٧	٦٤	١١	٣٤
٩٢	٢٠	٥٣	٤٧	٢٨	٨٠	٢٩	٨٠	٧٧	٣٧	٣٢	٦١	٣٧	٨٧	١٨
٥٨	٦٦	٠٩	٢١	٥٣	٥١	٥١	٦٥	٧١	٨٢	٧٩	٤٤	٩٢	٦٧	٩٤
٢٧	٣٢	٩١	٩٦	٥١	٣٥	١٤	٩٨	٦٥	٢٠	١٤	٢١	٤٤	٧٣	٤٠
٣١	٧٠	٠٦	٣٠	٣١	٥٣	٥٢	٧٩	٢٧	٥٠	٩١	٤٨	١٤	٧١	٦٣
٤٠	٨٢	٠٠	١٥	٩٥	١١	٥٠	٨٤	٨١	١٤	٩٦	١٨	١٥	٧٠	٣٩
٥٩	٦٥	٥٨	٠٦	٤٦	٣٢	٨٧	٢٣	٥٤	٧٧	٥٤	١٠	٢٥	٩١	٥٩
٦٤	٦١	٣٠	٢٤	٨٤	١٩	٩٧	٩٦	٣٧	٤٩	٥٢	٦٨	٥٣	٢٨	٠٥
٤٣	٥٦	٤٢	٥٥	٢١	٧٤	٢٨	١٠	٥٤	٠٢	٣٥	٧٠	٨٢	١٢	٦٣
٠٩	٨٢	٢٦	٤٠	٧٠	٧٠	٧٠	٨٨	١٣	١٦	٠٥	١٣	٨٢	٤٦	٣٣
٦٨	٧٩	٤٠	٠٠	٨٣	٤٣	٦٥	٦٨	٨٩	٤٠	١٤	٥٥	٦٧	٦٠	٨٥
٣٦	٨٠	٩٢	١٩	٧٦	٢١	٤٧	٤٠	١٩	٢٥	٩٨	٧٦	٩٨	٥٥	٢٠
٠٧	٣٦	٢٤	٤٣	٠٣	٦٢	٢٥	٤٧	٥٣	١٢	٩٠	٦٦	٥١	٣٢	٦١
٧٢	٢٧	١٣	٩٥	٢٣	١٧	٣٩	٣٢	١٤	٥٨	٩٠	٥٣	٥٤	٧٧	٦٨

ለ1	፳፬	፬	፬	፳	፳፬	፳፬	11	፻፻	፳1	፳፱	፬፬	፻፻	፳፳	11	ለ፬
፶፳	፳፳	፳፬	ለ፳	፬	፻፱	፻፬	፻1	1፬	፳፻	፬፻	፻፻	፬፬	፳1	፳፳	
፶፳	፻፻	፻፬	፳፬	፳፬	፳፳	፳፳	፬፳	፳1	1፳	፻፳	፬፻	፳፬	፳፳	፳፳	ለ፬
፻፳	፶፬	፶፳	፬፻	፳፬	፬፶	፳፻	፳፬	፻፻	፳፳	፳፳	፳፬	፬፬	፻፳	፳፳	፳1
፳፬	፶፳	፳፬	ለ፳	፳፳	፶፻	1፬	፻፶	፳፬	ለ፳	ለ፳	፳፳	፳፳	፳፳	፳፬	11

፳፳	1፱	፳፳	፳፳	ለ፳	ለ፶	፳፳	፳፬	፳፶	፳፳	፬፬	፳፻	፬፻	፳፳	፳፳	፳፬
፬፳	፬፻	፶፬	፶፶	፳፳	፳፳	፳፳	፳፳	፳፳	፳፬	ለ፬	፳፳	፳፳	፬፶	፬፶	ለ፳
፳፳	፬1	፶፳	፳፬	፳፻	1፬	ለ፬	፬፬	፻፳	፳፶	፻፬	፳፬	፳፬	1፶	1፻	፳፬
፳፬	፬፳	1፬	፻፳	፬፶	፳፳	1፳	፶፬	፳፳	፶1	፬፳	፬፳	፶፳	፳፳	፳፳	፻፬
፳1	፻፱	1፻	፶፶	፻፻	፳፳	፶፳	፶፳	1፳	፳፬	ለ፬	፶፳	፳፬	፳፬	፳፬	፳፳

جدول مناحات المنحنى المعتدل

٠,٠٩	٠,٠٨	٠,٠٧	٠,٠٦	٠,٠٥	٠,٠٤	٠,٠٣	٠,٠٢	٠,٠١	٠,٠٠	٢٠
٠,٣٥٩	٠,٣١٩	٠,٢٧٩	٠,٢٣٩	٠,١٩٩	٠,١٥٩	٠,١٢٠	٠,٠٨٠	٠,٠٤٠	٠,٠٠٠	٠,٠
٠,٠٧٥٣	٠,٠٧١٤	٠,٠٦٧٥	٠,٠٦٣٦	٠,٠٥٩٦	٠,٠٥٥٧	٠,٠٥١٧	٠,٠٤٧٨	٠,٠٤٣٨	٠,٠٣٩٨	٠,١
٠,١١٤١	٠,١١٠٣	٠,١٠٦٤	٠,١٠٢٦	٠,٠٩٨٧	٠,٠٩٤٨	٠,٠٩١٠	٠,٠٨٧١	٠,٠٨٣٢	٠,٠٧٩٣	٠,٢
٠,١٥١٧	٠,١٤٨٠	٠,١٤٤٣	٠,١٤٠٦	٠,١٣٦٨	٠,١٣٣١	٠,١٢٩٣	٠,١٢٥٥	٠,١٢١٧	٠,١١٧٩	٠,٣
٠,١٨٧٩	٠,١٨٤٤	٠,١٨٠٨	٠,١٧٧٢	٠,١٧٣٦	٠,١٧٠٠	٠,١٦٦٤	٠,١٦٢٨	٠,١٥٩١	٠,١٥٥٤	٠,٤
٠,٢٢٢٤	٠,٢١٩٠	٠,٢١٥٧	٠,٢١٢٣	٠,٢٠٨٨	٠,٢٠٥٤	٠,٢٠١٩	٠,١٩٨٥	٠,١٩٥٠	٠,١٩١٥	٠,٥
٠,٢٥٥٩	٠,٢٥١٨	٠,٢٤٨٦	٠,٢٤٥٤	٠,٢٤٢٢	٠,٢٣٨٩	٠,٢٣٥٧	٠,٢٣٢٤	٠,٢٢٩١	٠,٢٢٥٧	٠,٦
٠,٢٨٥٢	٠,٢٨٢٣	٠,٢٧٩٤	٠,٢٧٦٤	٠,٢٧٣٤	٠,٢٧٠٤	٠,٢٦٧٣	٠,٢٦٤٢	٠,٢٦١٢	٠,٢٥٨٠	٠,٧
٠,٣١٣٣	٠,٣١٠٦	٠,٣٠٧٨	٠,٣٠٥١	٠,٣٠٢٣	٠,٢٩٩٥	٠,٢٩٦٧	٠,٢٩٣٩	٠,٢٩١٠	٠,٢٨٨١	٠,٨
٠,٣٣٨٩	٠,٣٣٦٥	٠,٣٣٤٠	٠,٣٣١٥	٠,٣٢٨٩	٠,٣٢٦٤	٠,٣٢٣٨	٠,٣٢١٢	٠,٣١٨٦	٠,٣١٥٩	٠,٩
٠,٣٦٢١	٠,٣٥٩٩	٠,٣٥٧٧	٠,٣٥٥٤	٠,٣٥٣١	٠,٣٥٠٨	٠,٣٤٨٥	٠,٣٤٦١	٠,٣٤٣٨	٠,٣٤١٣	١,٠
٠,٣٨٣٠	٠,٣٨١٠	٠,٣٧٩٠	٠,٣٧٧٠	٠,٣٧٤٩	٠,٣٧٢٩	٠,٣٧١٨	٠,٣٦٨٦	٠,٣٦٦٥	٠,٣٦٤٣	١,١
٠,٤٠١٥	٠,٣٩٩٧	٠,٣٩٨٠	٠,٣٩٦٢	٠,٣٩٤٤	٠,٣٩٢٥	٠,٣٩٠٧	٠,٣٨٨٨	٠,٣٨٦٩	٠,٣٨٤٩	١,٢
٠,٤١٧٧	٠,٤١٦٢	٠,٤١٤٧	٠,٤١٣١	٠,٤١١٥	٠,٤٠٩٩	٠,٤٠٨٣	٠,٤٠٦٦	٠,٤٠٤٩	٠,٤٠٣٢	١,٣
٠,٤٣١٩	٠,٤٣٠٦	٠,٤٢٩٢	٠,٤٢٧٩	٠,٤٢٦٥	٠,٤٢٥١	٠,٤٢٣٦	٠,٤٢٢٢	٠,٤٢٠٧	٠,٤١٩٢	١,٤
٠,٤٤٤١	٠,٤٤٣٠	٠,٤٤١٨	٠,٤٤٠٦	٠,٤٣٩٤	٠,٤٣٨٢	٠,٤٣٧٠	٠,٤٣٥٧	٠,٤٣٤٥	٠,٤٣٣٢	١,٥
٠,٤٥٥٥	٠,٤٥٣٥	٠,٤٥٢٥	٠,٤٥١٥	٠,٤٥٠٥	٠,٤٤٩٥	٠,٤٤٨٥	٠,٤٤٧٤	٠,٤٤٦٣	٠,٤٤٥٢	١,٦
٠,٤٦٣٣	٠,٤٦٢٥	٠,٤٦١٦	٠,٤٦٠٨	٠,٤٥٩٩	٠,٤٥٩١	٠,٤٥٨٢	٠,٤٥٧٣	٠,٤٥٦٤	٠,٤٥٥٤	١,٧
٠,٤٧٠٦	٠,٤٦٩٩	٠,٤٦٩٣	٠,٤٦٨٦	٠,٤٦٧٨	٠,٤٦٧١	٠,٤٦٦٤	٠,٤٦٥٦	٠,٤٦٤٩	٠,٤٦٤١	١,٨
٠,٤٧٦٧	٠,٤٧٦٢	٠,٤٧٥٨	٠,٤٧٥٠	٠,٤٧٤٤	٠,٤٧٣٨	٠,٤٧٣٢	٠,٤٧٢٦	٠,٤٧١٩	٠,٤٧١٣	١,٩
٠,٤٨١٧	٠,٤٨١٢	٠,٤٨٠٨	٠,٤٨٠٣	٠,٤٧٩٨	٠,٤٧٩٣	٠,٤٧٨٨	٠,٤٧٨٣	٠,٤٧٧٨	٠,٤٧٧٣	٢,٠
٠,٤٨٥٧	٠,٤٨٥٤	٠,٤٨٥٠	٠,٤٨٤٦	٠,٤٨٤٢	٠,٤٨٣٨	٠,٤٨٣٤	٠,٤٨٣٠	٠,٤٨٢٦	٠,٤٨٢١	٢,١
٠,٤٨٩٠	٠,٤٨٨٧	٠,٤٨٨٤	٠,٤٨٨١	٠,٤٨٧٨	٠,٤٨٧٥	٠,٤٨٧١	٠,٤٨٦٨	٠,٤٨٦٥	٠,٤٨٦١	٢,٢
٠,٤٩١٦	٠,٤٩١٣	٠,٤٩١١	٠,٤٩٠٩	٠,٤٩٠٦	٠,٤٩٠٤	٠,٤٩٠١	٠,٤٨٩٨	٠,٤٨٩٦	٠,٤٨٩٣	٢,٣
٠,٤٩٣٦	٠,٤٩٣٤	٠,٤٩٣٢	٠,٤٩٣١	٠,٤٩٢٩	٠,٤٩٢٧	٠,٤٩٢٥	٠,٤٩٢٢	٠,٤٩٢٠	٠,٤٩١٨	٢,٤
٠,٤٩٥٢	٠,٤٩٥١	٠,٤٩٤٩	٠,٤٩٤٨	٠,٤٩٤٦	٠,٤٩٤٥	٠,٤٩٤٣	٠,٤٩٤١	٠,٤٩٤٠	٠,٤٩٣٨	٢,٥
٠,٤٩٦٤	٠,٤٩٦٣	٠,٤٩٦٢	٠,٤٩٦١	٠,٤٩٦٠	٠,٤٩٥٩	٠,٤٩٥٧	٠,٤٩٥٦	٠,٤٩٥٥	٠,٤٩٥٣	٢,٦
٠,٤٩٧٤	٠,٤٩٧٣	٠,٤٩٧٢	٠,٤٩٧١	٠,٤٩٧٠	٠,٤٩٦٩	٠,٤٩٦٨	٠,٤٩٦٧	٠,٤٩٦٦	٠,٤٩٦٥	٢,٧
٠,٤٩٨١	٠,٤٩٨٠	٠,٤٩٨٠	٠,٤٩٧٩	٠,٤٩٧٨	٠,٤٩٧٧	٠,٤٩٧٧	٠,٤٩٧٦	٠,٤٩٧٥	٠,٤٩٧٤	٢,٨
٠,٤٩٨٦	٠,٤٩٨٦	٠,٤٩٨٥	٠,٤٩٨٥	٠,٤٩٨٤	٠,٤٩٨٤	٠,٤٩٨٤	٠,٤٩٨٣	٠,٤٩٨٢	٠,٤٩٨١	٢,٩
٠,٤٩٩٠	٠,٤٩٨٩	٠,٤٩٨٩	٠,٤٩٨٩	٠,٤٩٨٨	٠,٤٩٨٨	٠,٤٩٨٨	٠,٤٩٨٧	٠,٤٩٨٧	٠,٤٩٨٦٥	٣,٠

جدول ٢

احتمال الحصول على قيمة χ^2 المبنية بالجدول بطريق المصادفة							درجات الحرية
٠,٩٩	٠,٩٠	٠,٥٠	٠,١٠	٠,٠٥	٠,٠١	٠,٠٠١	
٠,٠٠٠١٥٧	٠,٠١٥٨	٠,٤٥٥	٢,٧٠٦	٣,٨٤١	٦,٦٣٥	١٠,٨٢٧	١
٠,٠٢٠١	٠,٢١١	١,٣٨٦	٤,٦٠٥	٥,٩٩١	٩,٢١٠	١٣,٨١٥	٢
٠,١١٥	٠,٥٨٤	٢,٣٦٦	٦,٢٥١	٧,٨١٥	١١,٣٤١	١٦,٢٦٨	٣
٠,٢٩٧	١,٠٦٤	٣,١٥٧	٧,٧٧٩	٩,٤٨٨	١٣,٢٧٧	١٨,٤٦٥	٤
٠,٥٥٤	١,٦١٠	٤,٣٥١	٩,٢٣٦	١١,٠٧٠	١٥,٠٨٦	٢٠,٥١٧	٥
٠,٨٧٢	٢,٢٠٤	٥,٣٤٨	١٠,٦٤٥	١٢,٥٩٢	١٦,٨١٢	٢٢,٤٥٧	٦
١,٢٣٩	٢,٨٣٣	٦,٣٤٦	١٢,٠١٧	١٤,٠٦٧	١٨,٤٧٥	٢٤,٣٢٢	٧
١,٦٤٦	٣,٤٩٠	٧,٣٤٤	١٣,٣٦٢	١٥,٠٠٧	٢٠,٠٩٠	٢٦,١٢٥	٨
٢,٠٨٨	٤,١٦٨	٨,٣٤٣	١٤,٦٨٤	١٦,٩١٩	٢١,٦٦٦	٢٧,٨٧٧	٩
٢,٥٥٨	٤,٨٦٥	٩,٣٤٢	١٥,٩٨٧	١٨,٣٠٧	٢٣,٢٠٩	٢٩,٥٨٨	١٠
٣,٠٥٣	٥,٥٧٨	١٠,٣٤١	١٧,٢٧٥	١٩,٦٧٥	٢٤,٧٢٥	٣١,٢٦٤	١١
٣,٥٧١	٦,٣٠٤	١١,٣٤٠	١٨,٥٤٩	٢١,٠٢٦	٢٦,٢١٧	٣٢,٩٠٩	١٢
٤,١٠٧	٧,٠٢٤	١٢,٣٤٠	١٩,٨١٢	٢٢,٣٦٢	٢٧,٦٨٨	٣٤,٥٢٨	١٣
٤,٦٦٠	٧,٧٩٠	١٣,٣٣٩	٢١,٠٦٤	٢٣,٦٨٥	٢٩,١٤١	٣٦,١٣٣	١٤
٥,٢٢٩	٨,٥٤٧	١٤,٣٣٩	٢٢,٣٠٧	٢٤,٩٩٦	٣٠,٥٧٨	٣٧,٦٩٧	١٥
٥,٨١٢	٩,٣١٢	١٥,٣٣٨	٢٣,٥٤٢	٢٦,٢٩٦	٣٢,٠٠٠	٣٩,٢٥٢	١٦
٦,٤٠٨	١٠,٠٨٥	١٦,٣٣٨	٢٤,٧٦٩	٢٧,٥٨٧	٣٣,٤٠٩	٤٠,٧٩٠	١٧
٧,٠١٥	١٠,٨٦٥	١٧,٣٣٨	٢٥,٩٨٩	٢٨,٨٦٩	٣٤,٨٠٥	٤٢,٣١٢	١٨
٧,٦٣٣	١١,٦٥١	١٨,٣٣٨	٢٧,٢٠٤	٣٠,١٤٤	٣٦,١٩١	٤٣,٨٢٠	١٩
٨,٢٦٠	١٢,٤٤٣	١٩,٣٣٧	٢٨,٤١٢	٣١,٤١٠	٣٧,٥٦٦	٤٥,٣١٥	٢٠
٨,٨٩٧	١٣,٢٤٠	٢٠,٣٣٧	٢٩,٦١٥	٣٢,٦٧١	٣٨,٩٣٢	٤٦,٧٩٧	٢١
٩,٥٤٢	١٤,٠٤١	٢١,٣٣٢	٣٠,٨١٣	٣٣,٩٢٤	٤٠,٢٨٩	٤٨,٢٦٨	٢٢
١٠,١٩٦	١٤,٨٤٨	٢٢,٣٣٧	٣٢,٠٠٧	٣٥,١٧٢	٤١,٦٣٨	٤٩,٧٢٨	٢٣
١٠,٨٥٦	١٥,٦٥٩	٢٣,٣٣٧	٣٣,١٩٦	٣٦,٤١٥	٤٢,٩٨٠	٥١,١٧٩	٢٤
١١,٥٢٤	١٦,٤٧٣	٢٤,٣٣٧	٣٤,٣٨٢	٣٧,٦٥٢	٤٤,٣١٤	٥٢,٦٢٠	٢٥

فهرست

٧	طبيعة علم الاحصاء	—	الفصل الاول
٢٦	جمع المعلومات	—	الفصل الثاني
١٩٢	التصنيف والتبويب	—	الفصل الثالث
١٦٢	التوضيح البياني	—	الفصل الرابع
٢٢٧	المتوسطات	—	الفصل الخامس
٢٦٥	التشتت	—	الفصل السادس
٢٨٥	الالتواء	—	الفصل السابع
٢٩٥	توفيق المنحنى المعتدل	—	الفصل الثامن
٣١٣	اختبار كاي تربيع	—	الفصل التاسع
٣٤٤	العينات	—	الفصل العاشر
٣٨٧	الارتباط	—	الفصل الحادي عشر
٤٢٠	الانحدار	—	الفصل الثاني عشر
٤٥٤	الارقام القياسية	—	الفصل الثالث عشر
٤٨٣	السلاسل الزمنية	—	الفصل الرابع عشر